62. (amended) A method of determining the progression of a tumor to a more malignant phenotype comprising measuring the level of amplification of a MAC117 gene in a tissue or tumor sample containing cells or increased expression of a MAC117 gene in [the tumor] a body sample, [the presence of] an increase in the amplification or increased expression of said MAC117 gene indicating a more malignant phenotype.

REMARKS

Claims 44, 46, 47 and 60-62 are pending in this application. Claim 45 has been cancelled without prejudice. Claims 44, 46 and 60-62 have been amended to more particularly define the invention. The specification has been amended for clarity in accordance with the suggestions in the Office Action and to correct some inadvertent typographical errors not addressed in the Office Action. All of the amendments to the specification are of a clarifying nature only and are supported as set forth below. Therefore, no new matter has been added by these amendments. In light of these amendments and the following remarks, applicants respectfully request reconsideration and allowance of the pending claims.

Applicants note that the Office Action states that claims 47, 61 and 62 are allowable over the prior art.

Applicants also note that the subject application was filed on October 21, 1987 and that the numerous final rejections have been withdrawn only to have new grounds of rejection

entered. Applicants believe that in view of these amendments and the following remarks, the lengthy list of new issues raised in the October 4, 1994 Office Action has been addressed. However, if any issues remain, applicants invite the Examiner to call applicants' undersigned attorney in order to efficiently resolve these issues before this application reaches it eight year anniversary.

I. Rejection under 35 U.S.C. § 112, first paragraph

A. The specification is objected to and claims 44-47 and 60-62 are rejected under 35 U.S.C. § 112, first paragraph, as allegedly failing to adequately teach how to make and/or use the invention. Specifically, the Office Action states that the claims as presently worded direct the result of amplification or expression measurements as indicative of cancer that is "caused" by the measured effects and that the cause and effect relationship is not enabled by the disclosure. The Office Action further states that the cause and effect relationship between amplification or increased expression and cancer is speculative and therefore lacks enablement regarding the "caused by" limitation in the claims, because other explanations of the data are possible and reasonable.

Claims 44 and 60-62 have been amended to more specifically recite the invention by deleting the objectionable phrase "caused". Therefore, these rejections are rendered moot. Specifically, claim 44 has been amended to recite that the presence of amplification or increased expression of the MAC117 gene is indicative of the presence of cancer or a cancer with a more

malignant phenotype. Claim 60 has been amended to recite that the detection of amplification, rearrangement or increased expression of the MAC117 as well as the abnormal expression of the protein product of the MAC117 gene indicates the presence of human cancer in a patient. Claim 61 has been amended to recite that cancers from patients whose tissue or tumor cell samples show amplification or increased expression of the MAC117 gene can be classified as cancers which are correlated with amplification or increased expression of the MAC117 gene. Claim 62 has been amended to recite that the presence of an increase in the level of amplification or increased expression of the MAC117 gene indicates a more malignant phenotype. Claims 46 and 47 depend from claim 44 and are addressed by the amendment to claim 44. Support for these amendments is provided in the overall teachings of the specification and particularly on pages 19-21, 22a-22e, 23 and Table 1, wherein it is taught that amplification or increased expression of the MAC117 gene is consistently correlated with tumor formation *in vivo* and transformation *in vitro* and is never found as a condition in normal cells. Also, certain cancers are correlated with amplification or increased expression of specific oncogenes, allowing for their identification as tumors with more malignant phenotypes. Thus, as pointed out on page 23, lines 16-23, overexpression of certain oncogenes can contribute to the progression of such tumors and the five- to tenfold amplification of the v-erbB-related gene of the present invention in a mammary carcinoma indicates that increased expression of this gene may have provided a selective advantage to this tumor. Applicants believe these amendments further clarify the invention, thereby mooting this rejection. Applicants therefore respectfully request withdrawal of this rejection.

B. Claims 44-47 and 60-62 are rejected under 35 U.S.C. § 112, first paragraph, as the disclosure is allegedly enabling only for claims limited to methods where the sample is a tissue or tumor sample containing cells from the patient. Specifically, the Office Action states that it is noted that claim 60, part (a), practices detection in a tissue sample but that part (b) is not so limited in that it cites a "body sample." The Office Action also notes that claim 62 practices measurements in the tumor but does not clearly direct said measurements to cellular portions of said tumor. Applicants believe these rejections are not based on the skilled artisan being unable to measure increased expression in a body sample, but rather the rejection is based on the Examiner's position that applicants have not proven that increased expression can be measured in a cell sample.

The claims have been amended to add that amplification is measured in a sample containing cells. Thus, the rejection is moot for amplification. However, applicants contend that the MAC117 gene product can be detected in body samples other than cells.

To support this contention, applicants submit a Declaration by Dr. Kraus (an unsigned copy is included herein and the signature copy will be forthcoming), providing data disclosed in the publication submitted as Exhibit A: "Elevated soluble c-erbB-2 antigen levels in the serum and effusions of a proportion of breast cancer patients. " (Leitzel et al., (1992) **10**:1436-1443) which demonstrate the detection of the protein product of the MAC117 (c-erbB-2) gene in serum and effusions from cancer patients. Specifically, the Declaration states that an enzyme linked immunosorbent assay (ELISA) has been developed to detect the extracellular

domain of the c-erbB-2 oncogene product in serum and effusions from cancer patients. The assay data showed that sera from 12 of 53 patients with metastatic or locally advanced breast cancer, zero of 69 controls, one of 31 patients with ovarian cancer and two of 124 patients with other cancers had soluble c-erbB-2 levels greater than or equal to 5 U/ml. Also, two of five effusions from breast cancer patients had an elevated soluble c-erbB-2 antigen level, compared with zero of 17 effusions from patients with benign disease. These data establish evidence that the detection of increased expression of the MAC117 gene can in some cases be carried out by analyzing serum and effusion fluid as well as tissue or tumor cell samples and that the claim language reciting "a body sample" as applied to increased expression is not overly broad in its scope. On this basis, the recitation in claim 60, part (b) of a body sample is consistent with these data as they apply to the claimed invention and has been left in the claim. Claim 62 has been amended to recite a body sample instead of "the tumor" for the same reasons cited above in support of the language of claim 60. Applicants therefore respectfully request withdrawal of this rejection.

C. Claims 44-47 and 60-62 are rejected under 35 U.S.C. § 112, first paragraph, as the disclosure is allegedly enabling only for claims limited to cancer being indicated as caused by increased expression of the MAC117 gene. The Office Action specifically states that, as presently claimed, both gene amplification and increased expression indicate cancer presence as caused by either effect and that this is not supported by the transformation results described in the specification. The Office Action contends that in the disclosures in the specification, only increased expression of either mRNA or protein correlates to cancer type characteristics of cells being studied. The Office Action states that the lack of cancer type transformation resulting from

the introduction into cells of multiple gene copies of MAC117 in experiments employing the SV40/MAC117 expression vector indicates that the MAC117 gene amplification itself is not indicative of cancerous cell characteristics.

Applicants contend that the data set forth in the specification demonstrate that amplification of the MAC117 is indicative of a correlation with cancer. Specifically, amplification of the MAC117 is consistently correlated with the presence of cancer *in vivo*, as evidenced by the fact that amplification of the MAC117 gene has never been detected in normal tissues or cells in the absence of a cancerous condition. The example cited in the Office Action, wherein the presence of multiple copies of the MAC117 gene failed to result in transformation of the transfected cells, describes an artificial condition of amplification of the MAC117 gene *in vitro*, which is not representative of the demonstrated effects of amplification of the MAC117 gene *in vivo*. The lack of cancer type transformation resulting from SV40/MAC117 DNA as shown in Table 1 compared to the LTR mediated results may mean only that, under the control of the SV40 promoter, the MAC117 gene is not expressed beyond a threshold level necessary to induce transformation *in vitro*, irrespective of the level of amplification, whereas under the control of the LTR, this level is reached and transformation is induced. The Examiner himself has pointed out on page 4, lines 4-10 of the Office Action that these *in vitro* data are subject to a variety of interpretations. These data can hardly be only interpreted to negate the overwhelming amount of *in vivo* data demonstrating that amplification is consistently associated with cancerous conditions and are not detected in the absence of a cancerous condition, as has been addressed in previous responses to this application.

Applicants' data demonstrate that increased expression of the MAC117 in the absence of amplification is also correlated with a cancerous condition in a significant number of cases. Thus, in a clinical assay, detecting the amplification of the MAC117 provides a rapid screening method for diagnosing a cancer, identifying a particular type of cancer or prognosing the survival rate of a patient diagnosed with cancer, because the detection of amplification alone would provide the necessary information and no further expression testing would be required. If the assay failed to detect amplification, the patient's sample could then be assayed for increased expression of mRNA or protein to confirm or rule out a cancerous condition, identify a particular type of cancer or prognose the survival rate of a patient diagnosed with a type of cancer which correlates with increased expression of the MAC117 gene in the absence of detectable amplification of the MAC117 DNA.

In summary, 1) both amplification and increased expression of the MAC117 DNA *in vivo* are consistently associated with a cancerous phenotype and are not observed in normal tissue and 2) the detection of either amplification or increased expression in a patient's body sample is consistently indicative of the presence of cancer, as claimed in the present invention. For these reasons, applicants respectfully request withdrawal of this rejection.

D. Claims 44-47 and 60-62 are rejected under 35 U.S.C. § 112, first paragraph as the disclosure is allegedly enabling only for claims limited to specific hybridization probes such as the insert in pMAC117 or the segment between Nco I and Acc I and that no antibody probes are enabled. The Office Action states that the reason for this rejection is the lack of instantly enabled

epidermal growth factor (EGF) receptor protein or nucleic acid sequence enablement. The Office Action contends that because probes must be prepared to distinguish MAC117 sequences and epitopes from very similar EGF receptor embodiments, the EGF receptor gene and protein are essential subject matter and that the instant disclosure does not include EGF receptor sequence information or protein epitope information. The Office Action goes on to state that a number of references of printed publications are cited regarding various EGF receptor disclosures and that, according to M.P.E.P. §§ 706.03 (n) and 706.03(z), reference to these printed publications is insufficient for the disclosure of essential material. The Office Action recites that applicants are required to amend the disclosure to include the material incorporated by reference and include a Declaration stating that the amendatory material consists of the same material incorporated by reference.

Applicants have amended the specification to included the material described by the Office Action as essential subject matter. Specifically, at page 8, the specification has been amended to include the nucleotide and amino acid sequence of the EGF receptor gene as disclosed in Ullrich *et al.*, *Nature* 309:418-425 (May 1984) and submitted as Exhibit B, which is cited as incorporated by reference in the specification at page 2, lines 21-22 and page 3, lines 15-16. Furthermore, in accordance with the proper incorporation by reference of this essential subject matter, included herewith is a Declaration by Dr. Matthias Kraus (an unsigned copy is included herein and the signature copy will be forthcoming), wherein he states that the amendatory material consists of the same material incorporated by reference in the specification at

the cites described. It is believed that this amendment and Declaration moot this rejection and applicants respectfully request its withdrawal.

II. Rejection under 35 U.S.C. § 112, second paragraph

Claims 44-47 and 60-62 are rejected under 35 U.S.C. § 112, second paragraph as allegedly being indefinite for failing to particularly point out and distinctly claim the subject matter regarded as the invention. Specifically, the Office Action addresses the following issues:

A. Claims 44-47 and 60-62 are allegedly vague and indefinite in that the metes and bounds of what applicants mean regarding the practice of the phrase "a MAC117 gene" as cited in claim 44, line 3 or "the MAC117 gene" as cited in claim 60, lines 4 and 8 are not clearly defined. The Office Action states that the specification is unclear regarding the definition of the MAC117 gene for two reasons: 1) The Office Action inquires as to what is meant on pages 4 and 9d, lines 2-5 and 10-12, respectively, wherein the gene is described as having the nucleotide sequence as shown in Fig. 1. The specific question is whether this description refers to the sequence shown at the bottom of Figure 2 and cited as 424 bases in length or whether this description refers to a gene containing at least the insert sequence of λMAC117 shown at the top of Figure 1 defined by a restriction map but not actually depicted as a detailed nucleotide sequence similar to that at the bottom of Figure 1. 2) The Office Action also states that claim 45 appears to limit said gene to comprising the Figure 1 sequence of 424 bases but since claim 45 depends from claim 44, this implies that claim 44 is broader in scope than claim 45 thus suggesting that what is meant by "a

MAC117 gene" is broader than only a gene containing the Figure 1 sequence of 424 bases. The Office Action requests clarification regarding the basis for defining the gene and what distinctness measurement defines the MAC117 metes and bounds.

The standard textbook definition of a gene is a DNA sequence that codes for amino acids. (Watson et al., Molecular Biology of the Gene, Fourth Edition. Benjamin/Cummings Publishing Company, Inc., 1987, p. 233). Thus, by their disclosure on page 20, lines 23-26, that Figure 1 shows two exons of the MAC117 gene which encode part of the polypeptide chain of the MAC117 gene product, applicants have defined the MAC117 gene. Specifically, the MAC117 gene is the gene encoding the MAC117 protein. Figure 1 shows only a portion of the MAC117 gene. Additional parameters by which the MAC117 gene is defined are set forth in the specification as follows. On page 5, under the description of Figure 2, it is described that "DNA (from normal human placenta, A431 cells or human mammary carcinoma MAC117, prepared as described on pages 10-11) was cleaved with Eco RI (as described on pages 11-12), separated by gel electrophoresis in agarose gels (as described on page 12) and transferred to nitrocellulose paper." (as described on page 12). These DNAs were then probed with v-erbB probe which was "...a mixture of the 0.5-kbp Bam HI-Bam HI fragment and 0.5-kbp Bam HI-Eco RI fragment of avian erythroblastosis proviral DNA." (page 6, lines 7-9). This probe was prepared as described on pages 13-14. This particular probe was used to identify any DNA in human mammary carcinoma MAC117 which would cross-hybridize with gene sequences from the v-erbB gene because "[r]ecent analysis of the v-erbB gene and the EGF receptor gene indicates that the v-erbB gene is a part of the EGF receptor gene and codes for the internal domain and transmembrane

portion of the receptor." (page 2, lines 16-19). The results of the application of the v-erbB probe to these DNAs showed that "...DNA prepared from tissue of a human mammary carcinoma, MAC117, showed a pattern of hybridization that differed both from that observed with DNA of normal human placenta and from that observed with the A431 squamous-cell carcinoma line, which contains amplified epidermal growth factor (EGF) receptor genes. In A431 DNA, four Eco RI fragments were detected that had increased signal intensities compared to those of corresponding fragments in placenta DNA (Fig. 2A). In contrast, MAC117 DNA contained a single 6-kilobase pair (kbp) fragment, which appeared to be amplified compared to corresponding fragments observed in both A431 and placenta DNA's (Fig. 2A). These findings indicate that the MAC117 tumor contained an amplified DNA sequence related to, but distinct from, the cellular erbB proto-oncogene." (page 19, lines 8-23).

After having identified a genetic region of MAC117 DNA which cross-hybridized with an EGF receptor gene sequence, but which was distinct from the EGF receptor gene sequence, applicants cloned the MAC117 DNA as described on pages 17-18. As a result of screening of the library generated as described, "[t]en of 14 hybridizing phages contained a 6-kbp Eco RI fragment. Figure 1 shows the physical map of one of these phages, λMAC117, and pMAC117, a pUC12 subclone containing a 2-kbp Bam HI fragment of λMAC 117 that hybridized with v-erbB probes. The region of pMAC117 to which v-erbB hybridized most intensely was flanked by Acc I and Nco I sites." (page 18, lines 18-25). "The nucleotide sequence of the portion of pMAC117 located between the Nco I and Acc I sites contained two regions of nucleotide sequence homologous to v-erbB separated by 122 nucleotides (Fig. 1).

These regions shared 69 percent nucleotide sequence identity with both the v-erbB and the human EGF receptor gene. The predicted amino acid sequence of these regions was 85 percent homologous to two regions that are contiguous in the EGF receptor sequence. Furthermore, these two putative coding regions of the MAC117 sequence were each flanked by the AG and GT dinucleotides that border the exons of eukaryotic genes. These findings suggest that the sequence shown in Fig. 1 represents two exons, separated by an intron of a gene related to the erbB/EGF receptor gene." (page 20, lines 12-26).

Applicants also "...assembled a full length normal human MAC117 clone from overlapping cDNA clones (Figs. 5A,B)." (page 22c, lines 5-6). "Fig. 5A shows the restriction map of complementary DNA of MAC117 encompassing the entire coding region of the gene. Clone pMAC137 was isolated from an oligo dT primed human fibroblast cDNA library using a 0.8-kbp Acc I fragment from the 3' terminus of pMAC117 as probe. Clones λMAC30, λMAC10' and λMAC14-1 were subsequently isolated from a randomly primed MCF-7 cDNA library using cDNA fragments as probes." (page 9a, lines 2-7). "Fig. 5B illustrates 3 probes, a, b, and c, representing the 5', a middle portion, and the entire coding region, respectively..." (amended to appear on page 9a, prior to amendment, page 22a, lines 10-11).

Thus, the MAC117 gene is a sequence of DNA which contains a 6-kbp Eco RI fragment which binds a probe derived from the EGF receptor gene but which can be distinguished from the EGF receptor gene because the same probe binds different Eco RI fragments in A431 cells which are known to have an amplified EGF receptor gene. The MAC117 gene also contains

a 6-kbp Eco RI fragment, representing two exons as shown in Fig. 1 and consists of the coding region having the restriction map of Fig. 5A and which is represented in its entirety by the clones pMAC137, λMAC30, λMAC10' and λMAC14-1. Therefore, according to this definition of the MAC117 gene, claim 44 is broader in scope than claim 45, as implied by the claim language.

For the reasons set forth above, applicants believe that the definition of the MAC117 gene has been defined and therefore respectfully request withdrawal of this rejection.

B. Claims 45 and 46 are allegedly vague and indefinite in that it is unclear what distinguishes the two claims. The Office Action states that claim 45 contains the phrase "comprises at least in part" whereas claim 46 only contains the word "comprises" instead of the above phrase, suggesting that because the claims have different wording, claims 45 and 46 should have different meaning. The Office Action requests clarification of what different meaning is meant regarding claims 45 and 46. The Office Action also states that the phrase "at least in part " is confusing in that "at least" suggests a minimal content and "in part" suggests that more is present than the "at least" portion.

Applicants contend that the scope of claim 45 includes any identifying part of the 424 nucleotide sequence listed and that the scope of claim 46 is distinguished as the 424 nucleotide sequence itself. However, to expedite prosecution of this application to issue, claim 45 has been cancelled without prejudice and claim 46 has been amended to more particularly describe the invention. Specifically, claim 46 has been amended to recite a gene that encodes a polypeptide

encoded by a gene which comprises the nucleotide sequence recited in claim 46. Thus, the nucleotide sequence listed can vary in accordance with the degeneracy of the genetic code such that the resulting polypeptide is the same regardless of the specific nucleotide sequence. Applicants therefore respectfully request that this rejection be withdrawn.

C. In claim 60, line 4, the phrase "the MAC117 gene" allegedly lacks antecedent basis for a singular form such as suggested by the wording "the" MAC117 gene.

Claim 60 has been amended to establish antecedent basis for the MAC117 gene, thereby mooting this rejection. Applicants therefore respectfully request withdrawal of this rejection.

D. Claim 44 allegedly cites the intended method to be directed to "diagnosing or evaluating" in line 1 but only accomplishes in recited steps what is deemed "diagnosing" in line 6 cited therein as "indicating the presence of cancer". The Office Action also states that the evaluation criteria which correlate the extent of cancer seriousness to MAC117 amplification or increased expression are not evident from the claim but only become possible interpretations after contemplating the claim wording at length. The Office Action requests clarification of claim wording regarding what evaluation is intended to be practiced in claim 44 and those claims dependent therefrom.

Claim 44 has been amended to recite the evaluation of a cancer as having a more malignant phenotype. Support for this claim is found, for example, in claim 62 and on page 23, lines 14-27. Thus, the claim has been amended to more particularly define the invention. Applicants believe that this amendment renders moot this rejection and respectfully request its withdrawal.

III. Rejection under 35 U.S.C. § 102(a)

Claims 44-46 and 60 are rejected under 35 U.S.C. § 102(a) as allegedly being anticipated by either Semba et al. or Yamamoto et al. Specifically, the Office Action states that Semba et al. discloses the amplification of c-erbB-2 in human adenocarcinoma of the salivary gland and that these results read on the diagnostic methods as instantly claimed for the carcinomas therein analyzed. The Office Action further states that Yamamoto et al. discloses c-erbB-2 amplification in cancer cells and suggests that such amplification is sometimes involved in the neoplastic process. The Office Action alleges that this is a conservative evaluation of diagnostic use of the detection of such amplification, based on data for cell samples that serve as the basis for the instant invention also and therefore are equally supportive of diagnostic methods and therefore read on the above rejected claims. The Office Action further states that Semba et al. and Yamamoto et al. were published less than one year prior to the parent application serial number 06/836,414, thus making a 102(a) rejection appropriate due to priority given to said parent application for the subject matter of the above rejected claims.

Applicants assert that a 102(a) rejection to the claimed invention is inappropriate because the present invention was conceived and reduced to practice prior to the above cited references. In support of this assertion, included herewith is the Declaration under 37 C.F.R. § 1.131 of Drs. King, Kraus and Aaronson (an unsigned copy is included herein and the signature copy will be forthcoming), wherein these co-inventors declare that prior to October, 1985 and January, 1986, they reduced the claimed invention to practice, as evidenced by the publication (included as Exhibit C) of King et al., *Science* 229:974-976 (September, 1985), which contained all the data which formed the basis of the claimed invention and the published paper. For this reason, applicants contend that this rejection has been rendered moot and respectfully request its withdrawal.

IV. Objections to the specification for informalities

A. In the specification on page 2, line 14, the word "receptorcomplex" appears to be two words run together. This typographical error has been amended to distinguish these two words.

B. In the specification on page 4, line 15, the word "complimentary" is misspelled in the context in which it is used. The spelling has been amended.

C. In the specification on pages 5-7, Figures 1 and 2 are described in a manner which is confusing when compared to the content of Figures 1 and 2. This confusion is the result

of an inadvertent mislabeling of the figures of the invention and the specification has been amended to accurately describe the contents of Figures 1 and 2 by rearranging the descriptions in the text and by amending all references to these figures throughout the specification. No new matter has been added by this amendment, which consists of text already present in the specification, with minor clarifying modifications.

D. In the specification on page 7, line 4, reference is made to Figure 1A, whereas there is no Figure 1A. This reference is to Figure 2A and the specification has been amended accordingly.

E. In the specification on page 9a, Figure 5 is incompletely described and the description set forth actually describes only Figure 5A. A description of Figure 5B is included in the specification on page 22a, lines 10-13 but is not connected with the Figure 5 description on page 9a. The Figure 5 description on page 9a has been amended to indicate that it is the description for Figure 5A. The specification has also been amended so that the Figure 5B description on page 22a has been moved to page 9a under the heading of the description of Figure 5 and all references to Figure 5A and/or 5B have been amended for consistency.

F. In the specification on page 9a, lines 13 and 18, Fig. 1B, probes a and b are cited. Similarly, on page 9b, line 14, Fig. 1B, probe b is cited. The disclosure does not contain a Figure 1B nor are probes a and b described in Figure 1. Pages 9a and 9b have been amended to indicate that these are references to Figure 5B.

G. In the specification on page 9c, Table 1 is awkwardly located between the Brief Description of Figure 9 on page 9b and Figure 10 on page 9d. The specification has been amended to place Table 1 after page 35, which is the last page of the specification.

H. In the specification on page 9c, next to last line, the concentration range is incorrectly stated. The specification has been amended to accurately describe the concentration range.

I. In the specification on page 12, line 24, "nirocellulose" is misspelled. This typographical error has been amended to reflect the correct spelling.

J. In the specification on page 13, lines 14 and 15, Fig. 1A and 1B are cited. This has been amended to cite Fig. 2A and 2B.

K. In the specification on page 9, line 2 and page 13, line 12, respectively, PE7 and pE7 are cited to designate the apparently identical probe. The specification has been amended to consistently designate the probe as pE7.

L. In the specification on page 14, line 2, the enzyme designation "Hi" appears to be incorrect. The specification has been amended to correct this typographical error.

M. In the specification on page 15, lines 15-16, the Acc I- Nco I region is cited from Fig. 2. The specification has been amended to cite that this region is from Fig. 1.

N. In the specification on page 15, line 20, the designation "P^{12}" appears incorrect. The specification has been amended to recite the isotope ^{32}P.

O. In the specification on page 17, line 11, "dithiothretol" appears to be misspelled. This typographical error has been amended to reflect the correct spelling.

P. In the specification on page 17, lines 20-21, the phrase "plaques containing approximately 15,000 plaques" is unclear. This phrase has been amended to read "plates containing approximately 15,000 plaques."

Q. In the specification on page 17, line 1, it is confusing as to what is meant by the phrase "Cloning of MAC117." The phrase has been amended to recite "Cloning of λMAC117."

R. In the specification on page 18, line 12, the phrase "the 6-kbp fragment" is confusing in that a specific fragment is suggested, although there is no citation of any specific 6-kbp fragment in the specification before said phrase. The specification has been amended by deleting this out of context reference to a specific 6-kbp fragment.

S. In the specification on page 18, lines 20 and 26, Fig. 2 is cited as showing a map of a phage. The specification has been amended to recite Fig. 1 as showing the phage map.

T. In the specification on page 19, starting at line 8, Fig. 1A is cited three times as showing a hybridization pattern. The specification has been amended at these three cites to refer to Fig. 2A.

U. In the specification on page 19, lines 12 and 16, the words "sequmous" and "framents" appear to be misspelled. The specification has been amended to reflect the correct spellings of these words.

V. In the specification on page 20, first paragraph, Figs 1A and 1B are discussed. The specification has been amended to indicate that these are references to Figs 2A and 2B.

W. In the specification on page 20, lines 15 and 24, Figs 3 and 2, respectively, are discussed as showing nucleic acid sequence information. The specification has been amended to indicate that these are references to Fig. 1.

X. In the specification on page 22a, lines 2 and 19-20, citations directed to Kraus et al. are insufficient. The specification has been amended to include the full citation.

Y. In the specification on pages 22c and 22d, lines 9 and 16, citations directed to DiFiore et al. are insufficient. The specification has been amended to include the full citation.

Z. In the specification on page 22a, lines 3-5, a 0.8-kbp Acc I DNA fragment is cited with reference to Figure 1 and the other end of the fragment is not defined. The specification has been amended at page 9, line 4 and page 22a, line 4, to recite that the other end of this fragment is the 3' terminus of the MAC117 clone, as implied in the text.

AA. In the specification on pages 22c and 22d, last line and line 8, respectively, Table I is described, whereas the only table shown is designated as Table 1. The specification has been amended to consistently recite Table 1.

BB. In the specification on pages 24 and 27, lines 2 of each, Figure 1 is cited as if it discloses detection information. The specification has been amended to reflect that these are references to Figure 2.

CC. In the specification on page 24, lines 2-5, the Bgl I to Bam HI fragment of pMAC117 is cited as being used to detect the gene and its mRNA. This is claimed to be inconsistent with the record in that said fragment has been cited as being used only for detection of the gene. The same probe, consisting of the Bgl I to Bam HI fragment of pMAC117 is employed in the Southern blot of Figure 2 to detect the MAC117 gene, as stated in the specification on page 20, lines 1-5 and in the Northern blot of Figure 4 to detect the mRNA

transcribed from the MAC117 gene, as stated in the specification on page 21, lines 16-17. The probe to which reference is made on page 21 is described as the Bgl I-Bam HI fragment in the figure legend of Figure 4, on page 9, lines 1-2. The specification has been amended at page 21, line 16 to clarify that the Bgl I-Bam HI probe is being used in the Northern blot described in Fig. 4.

The Office Action further states that "Contrary to the above assertion the mRNA is disclosed as being performed using either probes a or b of Figure 5..." There is nothing contrary about any of these assertions. In the experiments described in Figures 2 and 4, the probe used was the Bgl I to Bam HI fragment of pMAC117. In the experiments described in Figure 6, the probes used were those described in Figure 5B as probes a and b. These are different experiments that employed different probes and there is nothing inconsistent or contrary in the specification regarding their use. Thus, there is no error in the statement on page 24, lines 2-5, that "...the isolation and use of the Bgl I to Bam HI restriction fragment of pMAC117 to specifically detect the gene and its mRNA has been set forth."

DD. In the specification on page 24, line 28, the probe designation "C" appears to be incorrect in comparison to Figure 5B. The specification has been amended to properly designate the probe as "c".

EE. In the specification on pages 26 and 28, lines 16 and 7, respectively, claim 4 is discussed whereas claim 4 has been cancelled. The specification has been amended to describe the amino acid sequence without reference to claim 4.

FF. In the specification on page 28, line 8, claim 1 is cited whereas claim 1 has been cancelled. The specification has been amended to describe the nucleic acid sequence without reference to claim 1.

GG. In the specification on page 26, lines 16-18, the sentence contains typographical errors which confuse its meaning. These errors have been amended to clarify the meaning of the sentence.

HH. In the specification on page 28, line 6, the word "form" is used incorrectly. The specification has been amended to recite "from".

II. In the specification on page 28, line 20, the word "boservations" is misspelled. The specification has been amended to recite the correct spelling of observations.

JJ. In the specification on page 28, line 25, the citation directed to Slamon et al. is insufficient. The specification has been amended to include the entire reference.

KK. In the specification on pages 29 and 31, lines 10 and 6, respectively, Figure 1 is cited incorrectly. The specification has been amended to indicate that these are references to Figure 2.

LL. In the specification on page 30, line 16, Figure 11 is cited whereas there is no Figure 11. The specification has been amended to indicate that this is a reference to Figure 10.

MM. In the specification on page 31, line 5, "demonstated" is misspelled. The specification has been amended to reflect the correct spelling of demonstrated.

NN. In the specification on page 34, line 7, "instrucitons" is misspelled. The specification has been amended to reflect the correct spelling of instructions.

In addition, the following amendments have been made to the specification which were not addressed in the Office Action but have been deemed necessary as clarifications to the specification.

The last sentence on page 34 of the specification has been amended to add the word "setting" to further clarify the meaning of that sentence.

On page 22c, line 6, "(Fig. 5)" has been amended to "(Figs. 5A,B)."

On page 27, line 2, "Figures 1" has been amended to "Figures 2."

On page 27, line 5, "5A" has been amended to "5B."

On page 27, line 12, the phrase "or increased expression" has been inserted between "amplification" and "of."

On page 27, line 17, the phrase "or its mRNA transcript" has been inserted between "gene" and "are."

On page 27, line 19, "Figure 1" has been amended to "Figure 2."

Other amendments have been made to the specification to correct typographical errors that were not pointed out in the Office Action.

All of the above amendments to the specification are of a clarifying nature and add no new matter to the specification. Applicants therefore respectfully request removal of these objections.

Pursuant to the above amendments and remarks, reconsideration and allowance of the pending application is believed to be warranted. The Examiner is invited and encouraged to

directly contact the undersigned if such contact may enhance the efficient prosecution of this application to issue.

A check in the amount of $870.00 is enclosed with a three-month Request for Extension of Time. This amount is believed to be correct; however, the Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 14-0629.

Respectfully submitted,

David G. Perryman
Registration No. 33,438

NEEDLE & ROSENBERG, P.C.
Suite 1200, The Candler Building
127 Peachtree Street, N.E.
Atlanta, Georgia 30303-1811
(404) 688-0770

Certificate of Mailing

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to:

Commissioner of Patents
and Trademarks
Washington, D.C. 20231

on this 4 day of April, 1995.

David G. Perryman — 4-4-95

David G. Perryman — Date

39906_1.WPD

EXHIBIT A

Elevated Soluble c-erbB-2 Antigen Levels in the Serum and Effusions of a Proportion of Breast Cancer Patients

By Kim Leitzel, Yoshio Teramoto, Ellen Sampson, June Mauceri, Beatrice C. Langton, Laurence Demers, Edward Podczaski, Harold Harvey, Shawn Shambaugh, Gena Volas, Susan Weaver, and Allan Lipton

Purpose: **An enzyme-linked immunosorbent assay (ELISA) for the extracellular domain of the c-*erb*B-2 oncogene product was developed and evaluated to determine if soluble c-*erb*B-2 could be detected in the serum and effusions of cancer patients.**

Patients and Methods: **Sera from 208 previously untreated or progressing cancer patients and 69 healthy controls were assayed in a double-antibody sandwich ELISA that used two monoclonal antibodies to the native extracellular domain of the c-*erb*B-2 receptor. Fisher's exact test was used to analyze the statistical significance of the frequency of elevated serum c-*erb*B-2 levels. Immunoprecipitation and Western blotting were used to characterize further the c-*erb*B-2 immunoreactivity in the serum of four breast cancer patients.**

Results: **Sera from 12 of 53 patients (23%) with metastatic or locally advanced breast cancer, zero of 69 controls, one of 31 patients with ovarian cancer (3%), and two of 124 other cancer patients (2%) had soluble c-*erb*B-2 values ≥ to 5 U/mL. The number of breast cancer patients with elevated serum c-*erb*B-2 levels was significantly greater than that of the control group ($P < .0001$), the ovarian cancer group ($P < .03$), and the other cancers group ($P < .0001$). Also, two of five effusions (40%) from breast cancer patients had an elevated soluble c-*erb*B-2 antigen level, compared with zero of 17 effusions from patients with benign diseases. Western blotting of four sera from breast cancer patients with elevated serum c-*erb*B-2 antigen levels produced bands of aproximately 105 kD that seemed to correlate in intensity with increasing ELISA serum levels.**

Conclusion: **Serum c-*erb*B-2 levels are elevated in approximately one fourth of patients with locally advanced or metastatic breast cancer.**

J Clin Oncol 10:1436-1443.

THE *NEU* ONCOGENE was first identified as the gene responsible for the transformation of NIH 3T3 cells that were transfected with DNA from carcinogen-induced rat neuroblastomas.[1] The human homolog of the rat *neu* oncogene (the c-*erb*B-2 [HER-2/*neu*] protooncogene) was discovered as a result of its substantial DNA homology to the epidermal growth factor (EGF) receptor.[2] The c-*erb*B-2 protein is a 185 kD transmembrane glycoprotein that contains an extracellular ligand-binding domain and intracellular tyrosine kinase activity.[2] A 30 kD protein that binds to both the c-*erb*B-2 protein and the EGF receptor has been reported recently as a putative ligand for c-*erb*B-2.[3] From immunohistochemical analysis, c-*erb*B-2 protein staining is observed only in low levels in the epithelial cells of most organs in normal human adults, and at slightly higher levels in fetal tissue; the c-*erb*B-2 gene is not amplified in these tissues.[4]

Amplification of the c-*erb*B-2 gene and overexpression of the protein was discovered in a human mammary carcinoma cell line.[5] Subsequently c-*erb*B-2 gene amplification was reported in 30% of primary human breast cancers.[6] A direct correlation between c-*erb*B-2 gene amplification and protein overexpression in breast cancer is known to exist, although in approximately 10% of cases, protein overexpression occurs without gene amplification.[7-9] A survey of the literature shows that c-*erb*B-2 gene amplification or protein overexpression occurs in approximately 10% to 40% of primary human breast cancers,[7-18] with an overall average of 20% from a combined total of almost 3,000 patients.[14]

Numerous studies have been performed to evaluate the prognostic significance of c-*erb*B-2 amplification in breast cancer.[14] In lymph node-positive breast cancer, some reports indicated a direct correlation between c-*erb*B-2 amplification or overexpression and poor survival,[7-9] whereas others failed to establish this correlation.[14,15] In lymph node-negative breast cancer patients, the results were also inconclusive. Some studies reported that c-*erb*B-2 amplification or overexpression was an indicator of poor prognosis,[10-13] whereas other reports again failed to corroborate these findings.[7-9,14,15] A recent case-control investigation that involved a large group of lymph node-negative patients (704 patients) reported that c-*erb*B-2 amplification statistically was a significant predictor of poor prognosis and was indepen-

From the Departments of Medicine, Pathology, Obstetrics and Gynecology, and Center for Biostastistics and Epidemiology, The Pennsylvania State University College of Medicine, The Milton S. Hershey Medical Center, Hershey, PA; Triton Diagnostics Inc, and Berlex Biosciences, Inc, Alameda, CA.

Submitted December 18, 1991; accepted May 4, 1992.

This research was supported by grants from the Eleanor Naylor Dana Charitable Trust, New York, NY, and the George L. Laverty Trust, Dauphin Deposit Bank and Trust Company, Harrisburg, PA.

Address reprint requests to Allan Lipton, MD, Department of Medicine, Division of Oncology, M.S. Hershey Medical Center, Box 850, Hershey, PA 17033.

0732-183X/92/1009-0011$3.00/0

dent of and had greater predictive power than most of the currently used prognostic factors, except tumor size.[16]

Amplification or overexpression of c-*erb*B-2 also occurred in ovarian,[7,19] lung,[20] and gastric carcinoma,[21,22] and was correlated with a poor prognosis in all of these malignancies.

The existence of soluble, or shed, forms of cell-surface transmembrane receptors increasingly has become a common phenomena. Soluble forms of the insulin,[23] EGF,[24] interleukin-2 (IL-2),[25] and tumor necrosis factor[26] receptors have been demonstrated in cell culture supernatants or in human serum. The quantitation of soluble IL-2 receptor levels by enzyme-linked immunosorbent assay (ELISA) in human serum has been shown to correlate reliably with disease activity in autoimmune inflammatory disorders, transplantation rejection, and specific infectious disorders. It reflects tumor burden and response to therapy in human T lymphotropic retrovirus type I-associated adult T-cell leukemia and hairy cell leukemia.[25] A soluble immunoreactive form of the c-*erb*B-2 receptor has also been detected in the sera of nude mice, which bore SK-OV-3 ovarian cancer or MDA-MB-361 breast cancer xenografts. These antigen levels correlated with both overexpression of the c-*erb*B-2 protein and increased tumor volume.[27] Increased levels of soluble c-*erb*B-2 have also been reported in the serum of three of 12 patients with breast cancer.[28] In this investigation, we report the presence of elevated levels of soluble c-*erb*B-2 detected by ELISA in the serum of a much larger group of patients with advanced breast cancer, and the lack of elevated soluble c-*erb*B-2 serum levels in a wide variety of other malignancies.

PATIENTS AND METHODS

Subjects

Serum c-*erb*B-2 levels were measured in 208 cancer patients (57 men and 151 women) who ranged from 25 to 86 years of age. Only patients with recently diagnosed but untreated cancer, or patients who had received previous treatment but had progression of disease at the time of drawing blood were selected for inclusion in this study. All other medical problems and medications were recorded. The clinical protocol was approved by the institutional review board and informed consent was obtained before venipuncture. The c-*erb*B-2 levels of effusions from 45 cancer patients and 17 patients with benign diseases were also determined.

Serum c-*erb*B-2 levels were also measured in 69 healthy control individuals (27 men and 42 women) who ranged from 19 to 76 years of age.

Serum and Effusion Preparation

For serum preparation, blood was drawn by forearm venipuncture and allowed to clot at room temperature for at least 1 hour before centrifugation. The blood tubes were then centrifuged at 500 × *g* for 10 minutes at room temperature, and the supernatant serum was collected, aliquoted, and stored at −70°C. Effusions (ascites or pleural fluid) were aspirated into Vacutainer tubes (Terumo Medical Corp., Elkton, MD) with .5 mL of a 3.5% solution of sodium citrate per tube. The effusion tubes were first centrifuged at 500 × *g* for 10 minutes at 4°C, then the supernatant was again centrifuged at 4,400 × *g* for 10 minutes at 4°C. This supernatant was collected, aliquoted, and stored at −70°C.

Generation of Monoclonal Antibodies

The monoclonal antibodies (MoAbs) selected for use in the double-antibody sandwich ELISA were generated by immunization of Balb/c mice with the human SK-BR-3 breast cancer cell line (TAb 259) or immunization with NIH 3T3 cells that were transfected with the full-length c-*erb*B-2 cDNA (TAb 257) as previously described.[27] Both MoAbs were of the immunoglobulin G1 (IgG1) isotype. Both MoAbs were reactive with the c-*erb*B-2 extracellular domain, which was purified from the conditioned medium of Chinese hamster ovary (CHO) cells that were transfected with the cDNA for the extracellular domain of the c-*erb*B-2 gene.[27]

The MoAb 36B7 used in the Western blot was generated by immunization of Balb/c mice with a plasma membrane preparation of NIH 3T3 cells that were transfected with the full length c-*erb*B-2 cDNA (NIH 3T3-tr).

c-erbB-2 Extracellular Domain Radioimmunoprecipitation

The extracellular domain of c-*erb*B-2 was purified from CHO cells that were transfected with the c-*erb*B-2 extracellular domain cDNA as described.[27] Purified extracellular domain was radiolabeled by the Iodogen method. TAb 259 or 257 were diluted in phosphate-buffered saline (PBS) with 0.5% bovine serum albumin (BSA) and 100 μL of diluted MoAb was incubated with radiolabeled extracellular domain (50,000 cpm per tube) overnight at room temperature. Bound counts were precipitated with 1% goat antimouse Ig in 5% polyethylene glycol 6,000 and 100 μL of 10% normal mouse serum. Tubes were centrifuged at 3,000 rpm for 15 minutes, and radioactivity in the pellets was determined in a gamma counter.

Cell Lines

The human cell lines listed in Table 1 were obtained from the American Type Culture Collection (Rockville, MD) and were cultured and passaged as directed.

Cell lysates were prepared by adding extraction buffer (.01 mol/L sodium phosphate, .15 mol/L sodium chloride [NaCl], 1% Triton X-100, .5% sodium deoxycholate, 1% sodium dodecyl sulfate [SDS], and 1 mmol/L phenylmethylsulfonylfluoride [PMSF] at pH 7.2) to the cell pellet, vortexing, and incubating on ice for 15 minutes. The cell lysates then were centrifuged at 10,000 × *g* for 10 minutes at 2°C to 8°C, and the supernatant was removed for testing in the ELISA.

c-erbB-2 ELISA

c-*erb*B-2 levels in human samples were determined with a double-antibody sandwich ELISA. Microtiter plates were coated by the addition of 100 μL per well of capture MoAb TAb 259 (diluted to 1 μg/mL with Tris-EDTA-NaCl [TEN] buffer). The TEN buffer consisted of 0.05 mol/L Tris, 0.001 mol/L ethylenedi-

aminetetraacetic acid (EDTA), and 0.15 mol/L NaCl at pH 7.3. After a blocking step, the microtiter plates were then dried and stored at 4°C before use. On the day the ELISA was run, the plate was equilibrated to room temperature, and 40 μL of sample diluent (1.25% mouse serum in TEN buffer) was added to each well.

Triplicate samples of the c-*erb*B-2 calibrators (10 μL) or duplicate samples (10 μL) of the human serum or effusions were added to the wells. The c-*erb*B-2 calibrators that were used were the supernatant fluids from NIH 3T3-tr cells. Regarding the calibrator solutions, 1 U/mL was equal to 0.25 ng/mL of the more than 95% pure extracellular domain that was obtained from the transfected CHO cells.

Fifty microliters of MoAb TAb 257 conjugated to horseradish peroxidase (0.7 μg/mL, diluted in TEN buffer that contained 1% BSA and 0.05% Tween 20) was added to each well. The plate was tapped gently to mix the reagents and then was incubated for 2 hours at room temperature. The plate was aspirated and washed six times with TEN buffer that contained 0.05% Tween 20. The plate was incubated with a substrate solution of tetramethylbenzidine (TMB) and hydrogen peroxide for 10 minutes at room temperature. Substrate conversion was arrested by the addition of 50 μL per well of 1 mol/L phosphoric acid. Substrate absorbance was determined at a test wavelength of 450 nm and a reference wavelength of 630 nm by an EL 312 microplate reader (Bio-Tek Instruments, Inc, Winooski, VT). Methodologic interassay precision ranged from 6% to 8%.

CEA, CA 549, and CA 125

Serum levels of carcinoembryonic antigen (CEA) and CA 549 were determined with the Tandem-CEA and Tandem-R CA 549 assay kits (Hybritech, Inc, San Diego, CA). Serum CA 125 levels were determined with the CA 125 radioimmunoassay kit (Amersham, Chicago, IL).

Serum Immunoprecipitation and Western Blotting

Sera from four cancer patients with substantially elevated serum c-*erb*B-2 antigen levels and sera from two cancer patients with normal serum c-*erb*B-2 antigen levels (determined by ELISA) were analyzed by immunoprecipitation and Western blotting. For immunoprecipitation, 40 μL of cancer patient serum was added to 1.0 μg of TAb 259. Then 200 μL of PBS was added, and the solution was incubated overnight at 4°C and mixed constantly. After the addition of 200 μL of Avid-AL (Bioprobe International Inc, Tustin, CA), the solution was incubated for 1 hour at room temperature and was mixed constantly. The immunoprecipitate was pelleted by centrifugation at 14,000 × *g* for 30 seconds. The supernatant was aspirated, and the pellet was washed three times by the addition of 1 mL of PBS that contained 0.05% Tween 20 (PBS and Tween), vortexed, and centrifuged at 14,000 × *g* for 30 seconds.

For the Western blotting procedure, 60 μL of sample buffer (62.5 mmol/L Tris, 5% SDS, 10% glycerol, 0.001% bromphenol blue, and 2% dithiothreitol at pH 6.75) was added to the immunoprecipitated pellet and was boiled for 5 minutes. The solution then was loaded onto a 10% polyacrylamide gel, and electrophoresis was conducted at 60 mA under reducing conditions to minimize interference and distortion in the 100 to 200 kD region of the blot because of the intact IgG proteins. The proteins were transferred onto nitrocellulose (70 V for 2 hours at 4°C), and the blot was blocked with PBS-containing 5% nonfat milk with 0.05% Tween 20 for 10 minutes. After it was rinsed in PBS and Tween, the blot was incubated with MoAb 36B7 (diluted to 2.5 μg/mL in PBS-containing 1.0% BSA and 0.05% Tween 20 [PBS, BSA, and Tween]) for 1 hour at room temperature. The blot was rinsed again in PBS and Tween and was incubated with peroxidase-conjugated AffiniPure rabbit antimouse IgG plus IgM (heavy and light chain specific; Jackson Immunoresearch Laboratories, Inc, West Grove, PA) diluted 1:5,000 in PBS, BSA, and Tween. After it was rinsed, the blot was developed with TMB and hydrogen peroxide with membrane enhancer (Kirkegaard and Perry Laboratories, Gaithersburg, MD).

Statistical Methods

The Fisher's exact test was used to analyze the significance of the number of elevated serum c-*erb*B-2 levels in the breast cancer group compared with the control group, ovarian cancer group, or the other cancers group (breast and ovarian cancer not included). Pearson correlation analysis was used to identify any linear correlation between c-*erb*B-2, CEA, CA 549, and CA 125 serum levels. All statistical tests were conducted by the Statistical Analysis System,[29] and a *P* value of .05 or less was accepted as statistically significant.

RESULTS

c-erbB-2 Extracellular Domain Radioimmunoprecipitation

The MoAb TAb 259 selected for use as the capture antibody in the double-antibody sandwich ELISA specifically reacted with the recombinant extracellular domain of c-*erb*B-2 as determined by immunoprecipitation of the radiolabeled protein (Fig 1). Anti-EGF receptor MoAb or other MoAbs with irrelevant specificities did not react with radiolabeled c-*erb*B-2 extracellular domain (data not shown). A similar binding curve was

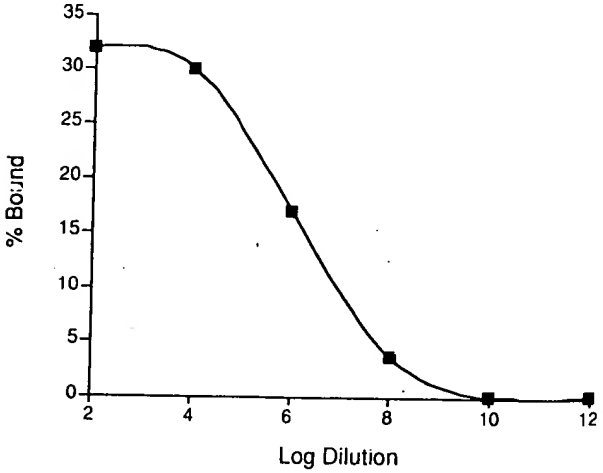


Fig 1. TAb 259 radioimmunoprecipitation of purified c-erbB-2 extracellular domain. Radiolabeled c-erbB-2 extracellular domain was incubated with dilutions of TAb 259 and precipitated with goat antimouse Ig. Percent bound was calculated as (specific bound counts minus nonspecific bound counts/total counts minus nonspecific bound counts) × 100.

obtained for TAb 257, the other MoAb used in the ELISA (data not shown). This combination of MoAbs previously was shown to produce excellent sandwich binding of immunoaffinity-purified extracellular domain from a panel of MoAbs generated to this domain of c-*erb*B-2.[27] These antibodies also immunoprecipitated the 185 kD c-*erb*B-2 protein from radiolabeled NIH 3T3-tr cells, but did not immunoprecipitate the EGF receptor from radiolabeled A431 cells.[27]

ELISA Detection of c-erbB-2 in Cell Lines

The double-antibody sandwich ELISA used in these experiments could detect reliably the presence of the c-*erb*B-2 protein above 3 U/mL (Fig 2). The specificity of the ELISA was determined by testing cell lysates from 12 human tumor cell lines and three normal human cell lines (Table 1). Cell lysates from those tumor cell lines that are known to overexpress the c-*erb*B-2 protein (SK-BR-3, MDA-361, and SK-OV-3) all showed substantial reactivity in the ELISA. Cell lysates of other cell lines tested did not react in the ELISA including A431, which overexpressed the EGF receptor, MCF-7, which has a high content of human milk fat globule protein, and SW 620, which overexpressed CEA (Table 1).

c-erbB-2 Levels in Cancer Patient Sera

The serum c-*erb*B-2 levels from 69 healthy control individuals were all ≤ 3.0 U/mL when tested in the ELISA (Fig 3). A critical step in the ELISA was the addition of mouse serum simultaneous with that of the patient or control serum samples. The mouse serum was added to eliminate a nonspecific binding activity that was observed in a small percentage (4%) of the 69 control sera when the ELISA was performed previously without the mouse serum (data not shown). In the group of 208 cancer patients, 12 of 53 patients (23%) with advanced breast cancer (51 with metastatic and two with locally advanced breast cancer) had c-*erb*B-2 levels above an arbitrary reference cutoff level of 5 U/mL (Fig 3).

One of 31 ovarian cancer patients (four with stage IV, 22 with stage III, one with stage II, and four with stage I disease) had a serum c-*erb*B-2 level more than 5 U/mL. This patient had stage III serous ovarian cancer and a

Table 1. c-erbB-2 Protein Content of Human Cell Line Lysates by ELISA

	Cell Line	Description	Reactivity	% C-erbB-2* Protein
Cancer cell lines	SK-BR-3	Breast adeno	High	.080
	MDA-361	Breast adeno	Medium	.010
	MDA-435P	Breast adeno	Negative	—
	MCF-7	Breast adeno	Negative (HMFG)	—
	SK-OV-3	Ovarian adeno	High	.080
	A431	Epidermoid CA	Negative (High EGFr)	—
	LS174T	Colon adeno	Negative (TAG 72)	—
	COLO 205	Colon adeno	Negative	—
	SW 620	Colon adeno	Negative (high CEA)	—
	HT-29	Colon adeno	Negative (high *ras*)	—
	SK-MES	Lung squamous CA	Negative	—
	HeLa	Cervical CA	Negative	—
Normal cell lines	HBL-100	Normal breast	Negative	—
	WI-38	Lung, diploid	Negative	—
	CCD18LU	Lung, diploid	Negative	—

Abbreviations: adeno, adenocarcinoma; CA, carcinoma.

*% c-erbB-2 protein = $\frac{\text{c-erbB-2 protein}}{\text{total protein}}$ in the cell lysate.

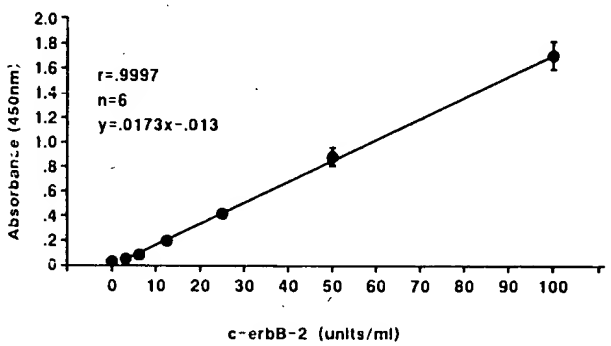


Fig 2. Calibration curve of c-erbB-2 ELISA. Data are from one ELISA. Data points are the mean ± SD of the absorbance at 450 nm of 3 to 4 wells per c-erbB-2 concentration. When no SD bars are visible, they are within the data point symbol; *r*, correlation coefficient.

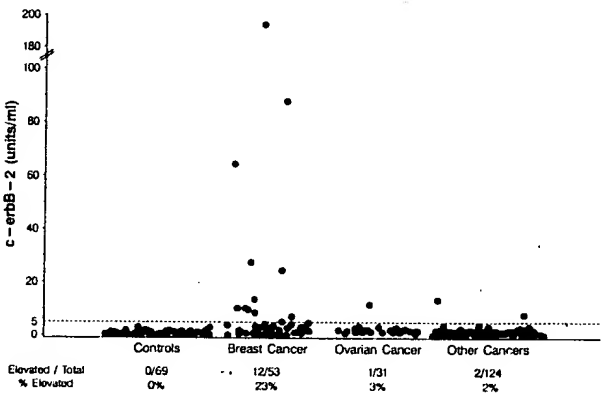


Fig 3. Serum c-erbB-2 levels in 69 controls and 208 cancer patients. The number of elevated c-erbB-2 serum levels in the breast cancer patient group was significantly greater than that of the control group ($P < .0001$), the ovarian cancer group ($P < .03$), and the other cancers group ($P < .0001$).

serum c-*erb*B-2 level of 11.6 U/mL. There was no correlation between c-*erb*B-2 and CA 125 serum levels in 23 ovarian cancer patients who had both values. Serum from a group of 124 patients with other types of cancer (breast and ovarian cancer not included), which was comprised of 75 cancers of epithelial origin, 29 sarcomas, nine melanomas, seven lymphomas, three leukemias, and one germ cell tumor (teratoma), were also tested in the ELISA. Only two of 124 of these patients (2%) had elevated c-*erb*B-2 values. One male patient had a stage I lymphocytic lymphoma of the right orbit; he was the only patient who had an elevated serum c-*erb*B-2 level (13.4 U/mL) of a group of seven lymphoma patients (14%). The other patient was a female with an immature teratoma (stage III) and a serum c-*erb*B-2 value of 7.8 U/mL.

The number of breast cancer patients with an elevated serum c-*erb*B-2 level was significantly greater than that of the control group ($P < .0001$), the ovarian cancer group ($P < .03$), and the other cancers group ($P < .0001$). Partial χ^2 analysis indicated that there was no significant difference, however, in the proportion of elevated c-*erb*B-2 serum levels between the control, other cancers, and ovarian cancer groups. A small number of breast cancer patients also had CEA and CA 549 measurements from the same serum sample that was used in the c-*erb*B-2 ELISA. There was no linear correlation between c-*erb*B-2 and CEA levels (n = 14), or between c-*erb*B-2 and CA 549 levels (n = 7).

*Immunoprecipitation and Western Blot of Elevated Serum c-*erb*B-2 Samples*

Sera from four cancer patients with elevated serum c-*erb*B-2 levels were subjected to immunoprecipitation and Western blotting in an attempt to determine the molecular weight of the immunoreactive antigen observed in the c-*erb*B-2 ELISA. MoAb 36B7 was selected for Western blotting because of its excellent reactivity under reducing conditions. MoAb 36B7 did not react with the EGF receptor as was determined by ELISA and Western blotting of A431 cell lysate, nor did it react in Western blots with conditioned medium from nontransfected NIH 3T3 cells (data not shown).

After immunoprecipitation and Western blotting, heterogenous bands at approximately 105 kD were observed for all four breast cancer patients with elevated serum c-*erb*B-2 ELISA levels (Fig 4, lanes 2, 3, 6, and 7). No bands were observed in two breast cancer patients without elevation of serum c-*erb*B-2 (Fig 4, lanes 4 and 5). A band of approximately 105 kD was also observed in the conditioned medium of NIH 3T3-tr cells (Fig 4, lane 1). The breast cancer patient serum with the highest serum c-*erb*B-2 level (194 U/mL) also seemed to have the most intense 105 kD band (Fig 4, lane 2).

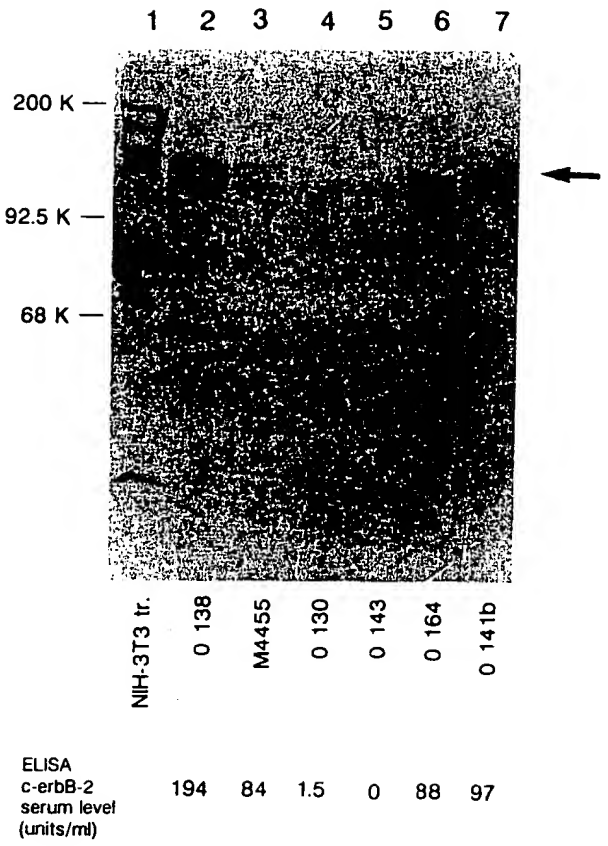


Fig 4. Western blotting of breast cancer patient serum. Samples were immunoprecipitated and Western blotted as previously described. Lane 1 is the conditioned medium of NIH 3T3 cells transfected with the full-length c-*erb*B-2 cDNA. Lanes 2 to 7 are the sera of six different breast cancer patients with ELISA serum c-*erb*B-2 levels shown below each lane.

*Effusion c-*erb*B-2 Levels*

The c-*erb*B-2 levels of effusions from 45 cancer patients and from 17 patients with benign diseases (nine with congestive heart failure, five with liver cirrhosis, and one each with peptic ulcer, uremia, and cardiomyopathy) were also determined in the ELISA (Fig 5). The benign effusions had a c-*erb*B-2 range of 0 to 3.9 U/mL. The 45 cancer patients with effusions included 39 with cancers of epithelial origin, three lymphomas, two melanomas, and one sarcoma. Four of the 45 cancer patients (9%) had c-*erb*B-2 effusion levels more than 3.9 U/mL (Fig 5). Of these four elevated effusions, two were from a total of five breast cancer patients (100, 53 U/mL), one of two gastric cancer patients (200 U/mL), and one of

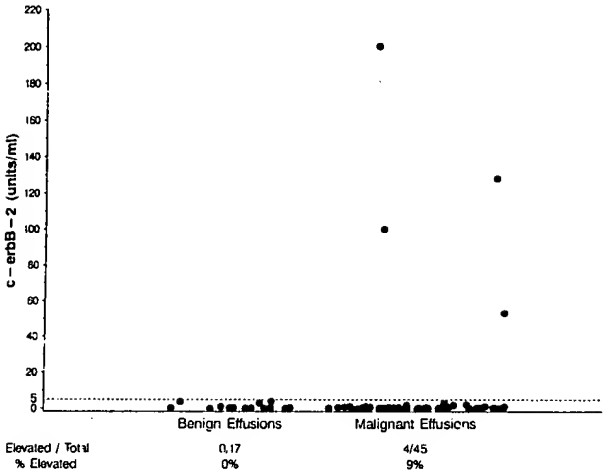


Fig 5. Soluble c-erbB-2 levels in 17 benign and 45 malignant effusions.

two cervical cancer patients (128 U/mL). None of 14 ovarian cancer patients (four stage IV, 10 stage III) had c-*erb*B-2 effusion levels more than 3.9 U/mL.

One of the two breast cancer patients who had elevated effusion c-*erb*B-2 levels (from the total of five breast cancer patients) had a serum sample drawn the day after the pleural fluid sample. The pleural fluid soluble c-*erb*B-2 level was 100 U/mL, and the serum sample was 88 U/mL. Western blotting of this patient's serum showed a band at 100 to 105 kD (Fig 4, lane 6). At the time of serum and pleural fluid collection, this patient had breast adenocarcinoma that had metastasized to soft tissue and bone; she progressed rapidly despite chemotherapy. Thus, in the only matched serum and effusion sample tested, there were soluble c-*erb*B-2 elevations in both fluids.

DISCUSSION

In this study, a double-antibody sandwich ELISA that used MoAbs that were reactive with the extracellular domain of the c-*erb*B-2 receptor demonstrated increased serum levels of a soluble form of the c-*erb*B-2 receptor in a portion (23%) of patients with advanced breast cancer. The serum levels of soluble c-*erb*B-2 were uniformly low in the control group of 69 healthy individuals. Elevated serum c-*erb*B-2 levels were observed in only one of 31 ovarian cancer patients (3%) and in only two of 124 patients (2%) with other cancers (breast and ovarian cancer not included). Also, two of five effusions from metastatic breast cancer patients (40%) had an elevated soluble c-*erb*B-2 level. Our results confirm and significantly expand the initial observation that a soluble derivative of the c-*erb*B-2 receptor can be found in the serum of a proportion of breast cancer patients.[28] Our results are also similar to the recent publication of Carney et al[30] who reported evaluations of soluble c-*erb*B-2 in the plasma of 24 of 105 patients (23%) with metastatic breast cancer. We further show that elevated serum c-*erb*B-2 rarely is found in other malignancies.

Immunoprecipitation and Western blot analysis of the elevated serum c-*erb*B-2 immunoreactivity observed in four breast cancer patients in our ELISA produced bands of approximately 105 kD. No bands were observed in two breast cancer patients with normal serum c-*erb*B-2 levels. This estimated molecular weight was similar to the 110 kD c-*erb*B-2 gene product-related protein[31] and the 105 kD extracellular domain of c-*erb*B-2[32] that have been reported in the conditioned medium of the SK-BR-3 human breast carcinoma cell line, and the 130 kD soluble c-erbB-2–related protein released from the BT474 human breast carcinoma cell line.[33] The 105 kD size of our c-*erb*B-2 immunoreactive protein in the serum of breast cancer patients also is close to the 97-115 kD band recently reported in the plasma of breast cancer patients,[30] and is also close to the predicted molecular weight of 118 kD calculated for the c-*erb*B-2 extracellular domain based on the gene sequence.[32]

The mechanism of release of soluble truncated c-*erb*B-2 receptor remains controversial. Soluble c-*erb*B-2 could be produced by posttranslational processing of the full-length c-*erb*B-2 gene, which includes cell-surface proteolysis that is similar to that suggested for the soluble IL-2 receptor[25] and the soluble TNF receptor.[26] Alternately, differential splicing and subsequent production of a soluble c-*erb*B-2 extracellular domain from a smaller transcript could occur. One study that used SK-BR-3 cells reported the expression of the full-length 5.0 kb c-*erb*B-2 transcript and a smaller 2.8 kb, which hybridized only to extracellular c-*erb*B-2 probes.[31] Another report failed to detect an alternate c-*erb*B-2 transcript and proposed that the soluble extracellular c-*erb*B-2 domain was released from the surface of SK-BR-3 cells by proteolysis, as supported by pulse-chase analysis.[32] In addition to proteolysis or alternative splicing, gene rearrangement could produce a truncated c-*erb*B-2 product, but no evidence for this has been shown to date.

The physiologic function of soluble extracellular c-*erb*B-2 is not known, but it could serve potentially as (1)an antagonist of normal cell-surface receptor-ligand interaction, (2) a carrier for the transportation, shielding, or destruction of c-*erb*B-2 ligand, or (3) a nonfunctional by product of normal proteolytic action. An

interesting proposal related to this third possibility is that proteolysis is likely to be important in the normal process of receptor down-regulation.[32] Also, as suggested, the remaining cell-associated cleavage product that is composed of the transmembrane and cytoplasmic domains could represent an oncogenic form of the c-*erb*B-2 receptor tyrosine kinase. Transformation could then occur by a proportional increase in the absolute number of truncated receptors caused by the proteolysis of overexpressed c-*erb*B-2 receptors on the tumor cell surface.

The lack of elevated serum c-*erb*B-2 in 30 of 31 ovarian cancer patients may have several explanations. First, our sample size may have been too small, and we may not have accrued enough patients with c-*erb*B-2–positive tumor cells to reach the 30% of ovarian cancer patients who have been shown to overexpress c-*erb*B-2.[7,19] The immunohistochemical staining of these ovarian tumors for c-*erb*B-2 addresses this question; these studies are in progress for the ovarian and breast tumors reported. Second, it is possible that the putative protease(s) that may release the extracellular domain of c-*erb*B-2 in breast cancer cells may not be present or active in ovarian tissue, or alternative splicing of the full length c-*erb*B-2 transcript may not occur.

There is a great need for more definitive serum markers for use in the detection, management, and response to treatment in breast cancer. Although other breast cancer serum markers such as CEA, CA 549, and CA 15-3 are sometimes used in clinical practice, they do not always reflect the extent of disease and prognosis in patients. A marker that would parallel more closely the clinical status of the patient clearly would be of great benefit. The critical role that the c-*erb*B-2 receptor plays in oncogenesis, combined with our results that show that 23% of patients with metastatic breast cancer have elevated serum levels of c-*erb*B-2, support the conclusion that serum c-*erb*B-2 levels deserve continued intensive study as a potential marker in breast cancer.

ACKNOWLEDGMENT

We thank Michelle Moore for her excellent help with the preparation of the manuscript. We are also grateful to the Outpatient Medicine Nursing Staff for blood collection.

REFERENCES

1. Shih C, Padhy LC, Murray M, et al: Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts. Nature 290:261-263, 1981

2. Maguire HC, Greene MI: The *neu* (c-*erb*B-2) oncogene. Semin Oncol 16:148-155, 1989

3. Lupu R, Colomer R, Zugmaier G, et al: Direct interaction of a ligand for the *erb*B2 oncogene product with the EGF receptor and p 185^{erbB2}. Science 249:1552-1555, 1990

4. Press MF, Cordon-Cardo C, Slamon DJ: Expression of the HER-2/*neu* proto-oncogene in normal human adult and fetal tissues. Oncogene 5:953-962, 1990

5. King CR, Kraus MH, Aaronson SA: Amplification of a novel v-*erb*B-related gene in a human mammary carcinoma. Science 229:974-976, 1985

6. Slamon DJ, Clark GM, Wong SG, et al: Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/*neu* oncogene. Science 235:177-182, 1987

7. Slamon DJ, Godolphin W, Jones LA, et al: Studies of the HER-2/*neu* proto-oncogene in human breast and ovarian cancer. Science 244:707-712, 1989

8. Tandon AK, Clark GM, Chamness GC, et al: HER-2/*neu* oncogene protein and prognosis in breast cancer. J Clin Oncol 7:1120-1128, 1989

9. Borg A, Tandon AK, Sigurdsson H, et al: HER-2/*neu* amplification predicts poor survival in node-positive breast cancer. Cancer Res 50:4332-4337, 1990

10. Wright C, Angus B, Nicholson S, et al: Expression of c-*erb*B-2 oncoprotein: A prognostic indicator in human breast cancer. Cancer Res 49:2087-2090, 1989

11. Ro J, El-Naggar A, Ro JY, et al: c-*erb*B-2 amplification in node-negative human breast cancer. Cancer Res 49:6941-6944, 1989

12. Paik S, Hazan R, Fisher ER, et al: Pathologic findings from the National Surgical Adjuvant Breast and Bowel Project: Prognostic significance of *erb*B-2 protein overexpression in primary breast cancer. J Clin Oncol 8:103-112, 1990

13. McCann A, Dervan PA, Johnston PA, et al: c-*erb*B-2 oncoprotein expression in primary human tumors. Cancer 65:88-92, 1990

14. Clark GM, McGuire WL: Follow-up study of HER-2/*neu* amplification in primary breast cancer. Cancer Res 51:944-948, 1991

15. Zhou DJ, Ahuja H, Cline MJ: Proto-oncogene abnormalities in human breast cancer: c-*erb*B-2 amplification does not correlate with recurrence of disease. Oncogene 4:105-108, 1989

16. Paterson MC, Dietrich KD, Danyluk J, et al: Correlation between c-*erb*B-2 amplification and risk of recurrent disease in node-negative breast cancer. Cancer Res 51:536-567, 1991

17. van de Vijver MJ, Peterse JL, Mooi WJ, et al: *Neu*-protein overexpression in breast cancer: Association with comedo-type ductal carcinoma in situ and limited prognostic value in stage II breast cancer. N Engl J Med 319:1239-1245, 1988

18. Iglehart JD, Kraus MH, Langton BC, et al: Increased *erb*B-2 gene copies and expression in multiple stages of breast cancer. Cancer Res 50:6701-6707, 1990

19. Berchuck A, Kamel A, Whitaker R, et al: Overexpression of HER-2/*neu* is associated with poor survival in advanced epithelial ovarian cancer. Cancer Res 50:4087-4091, 1990

20. Kern JA, Schwartz DA, Nordberg JE, et al: $p185^{neu}$ expression in human lung adenocarcinomas predicts shortened survival. Cancer Res 50:5184-5191, 1990

21. Park JB, Rhim JS, Park SC, et al: Amplification, overexpression, and rearrangement of the *erb*B-2 protooncogene in primary human stomach carcinomas. Cancer Res 49:6605-6609, 1989

22. Yonemura Y, Ninomiya I, Yamaguchi A: Evaluation of immunoreactivity for *erb*B-2 protein as a marker of poor short term prognosis in gastric cancer. Cancer Res 51:1034-1038, 1991

23. Gavin JG III, Buell DN, Roth J: Water-soluble insulin receptors from human lymphocytes. Science 178:168-169, 1972

24. Weber W, Gill GN, Spiess J: Production of an epidermal growth factor receptor-related protein. Science 224:294-297, 1984

25. Rubin LA, Nelson DL: The soluble interleukin-2 receptor: Biology, function, and clinical application. Ann Intern Med 113:619-627, 1990

26. Gatanaga T, Hwang C, Kohr W, et al: Purification and characterization of an inhibitor (soluble tumor necrosis factor receptor) for tumor necrosis factor and lymphotoxin obtained from the serum ultrafiltrates of human cancer patients. Proc Natl Acad Sci U S A 87:8781-8784, 1990

27. Langton BC, Crenshaw MC, Chao LA, et al: An antigen immunologically related to the external domain of gp 185 is shed from nude mouse tumors overexpressing the c-*erb*B-2 (HER-2/*neu*) oncogene. Cancer Res 51:2593-2598, 1991

28. Mori S, Mori Y, Mukaiyama T, et al: In vitro and in vivo release of soluble *erb*B-2 protein from human carcinoma cells. Jpn J Cancer Res 81:489-494, 1990

29. SAS Institute, Inc, SAS Procedures Guide, Version 6, ed 3, Cary, NC, SAS Institute, Inc, 1990, pp 705

30. Carney WP, Hamer PJ, Petit D, et al: Detection and quantitation of the human *neu* oncoprotein. J Tumor Marker Oncol 6:53-72, 1991

31. Alper O, Yamaguchi K, Hitomi J, et al: The presence of c-*erb*B-2 gene product-related protein in culture medium conditioned by breast cancer cell line SK-BR-3. Cell Growth and Diff 1:591-599, 1990

32. Zabrecky JR, Lam T, McKenzie SJ, et al: The extracellular domain of p185/*neu* is released from the surface of human breast carcinoma cells, SK-BR-3. J Biol Chem 266:1716-1720, 1991

33. Lin YJ, Clinton GM: A soluble protein related to the HER-2 proto-oncogene product is released from human breast carcinoma cells. Oncogene 6:639-643, 1991

This Page is inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- [x] BLACK BORDERS
- [ ] IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- [ ] FADED TEXT OR DRAWING
- [x] BLURED OR ILLEGIBLE TEXT OR DRAWING
- [ ] SKEWED/SLANTED IMAGES
- [ ] COLORED OR BLACK AND WHITE PHOTOGRAPHS
- [ ] GRAY SCALE DOCUMENTS
- [ ] LINES OR MARKS ON ORIGINAL DOCUMENT
- [ ] REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- [ ] OTHER: _____

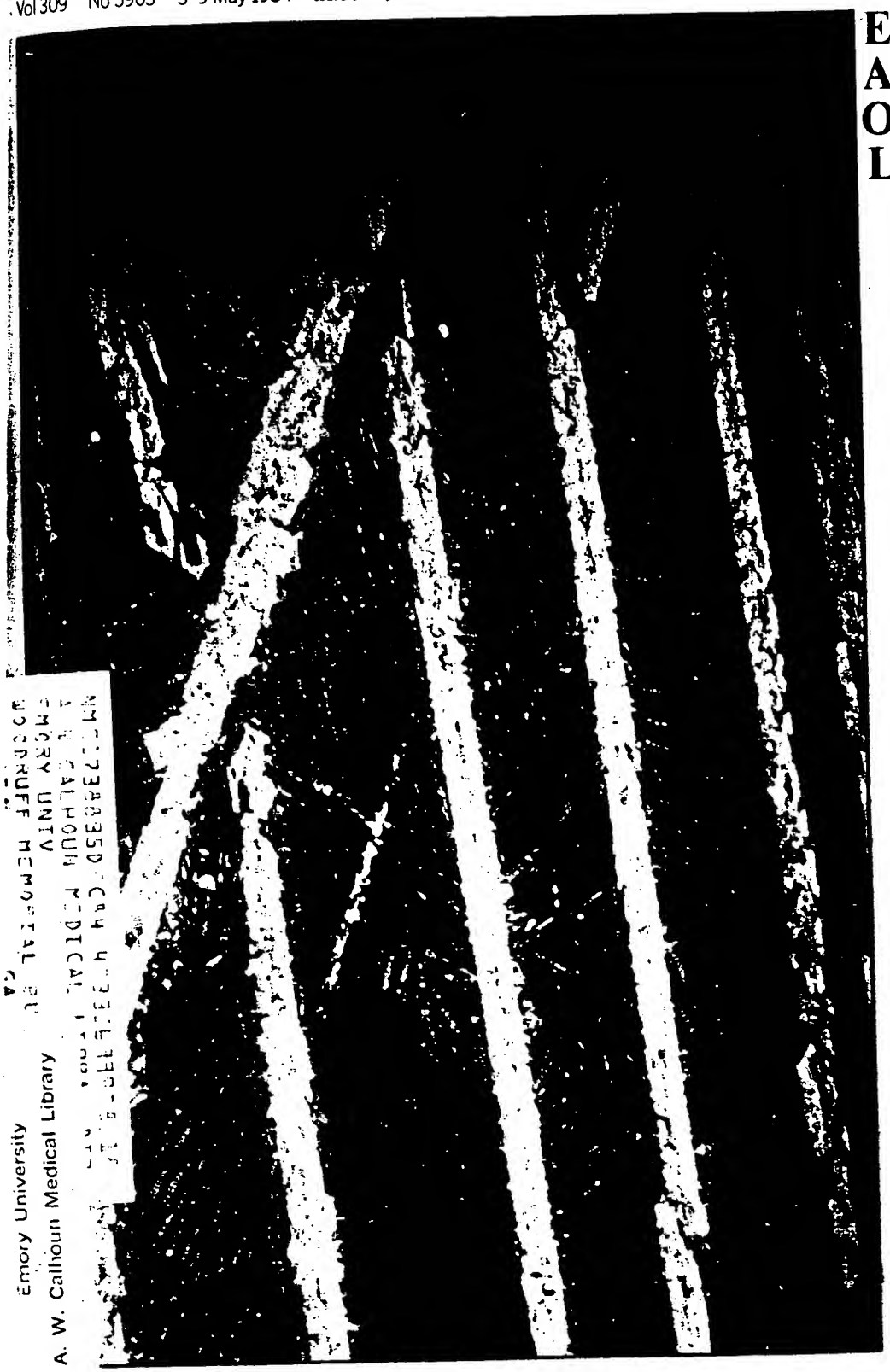
IMAGES ARE BEST AVAILABLE COPY.
As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox

EXHIBIT B

nature

INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

Vol 309 No 5963 3-9 May 1984 £1.80 $4.50



EMPLACEMENT AND COOLING OF KOMATIITE LAVAS

Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells

A. Ullrich, L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden*, T. A. Libermann*, J. Schlessinger*, J. Downward†, E. L. V. Mayes†, N. Whittle†, M. D. Waterfield† & P. H. Seeburg

Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA
* Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot, Israel
† Protein Chemistry Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

The complete 1,210-amino acid sequence of the human epidermal growth factor (EGF) receptor precursor, deduced from cDNA clones derived from placental and A431 carcinoma cells, reveals close similarity between the entire predicted v-erb-B mRNA oncogene product and the receptor transmembrane and cytoplasmic domains. A single transmembrane region of 23 amino acids separates the extracellular EGF binding and cytoplasmic domains. The receptor gene is amplified and apparently rearranged in A431 cells, generating a truncated 2.8-kilobase mRNA which encodes only the extracellular EGF binding domain.

EPIDERMAL growth factor (EGF), a potent mitogenic polypeptide, initiates cellular responses by binding to specific receptors on the surface of target cells (for reviews see refs 1, 2). Like several other growth factors, the binding of EGF to its cell surface receptor triggers a cascade of intracellular events, including induction of a tyrosine kinase activity intrinsic to the EGF receptor[1–5]. Shortly after binding, the EGF–receptor complexes are localized in clathrin-coated regions of the plasma membrane and are then internalized by the cell.[6,7].

How these events lead to the stimulation of DNA synthesis and cellular proliferation are unknown. The recent observation that the receptors for platelet-derived growth factor (PDGF)[8,9], insulin[10,11] and insulin-like growth factor (IGF-I)[12], are also tyrosine-specific protein kinases suggested that tyrosine phosphorylation might be an important early step in triggering proliferation. The tyrosine-specific protein kinase activity of several retroviral transforming proteins (reviewed in ref. 13) has led many workers to search for a connection between growth factor receptors and transforming proteins. In their analysis of partial protein sequences, Downward *et al.*[14] demonstrated close similarity between the EGF receptor and the avian erythroblastosis virus (AEV) v-*erb-B* transforming protein[15]; evidence was presented suggesting that the viral transforming protein contained only the transmembrane and tyrosine kinase domains of the avian EGF receptor, and lacked most of the extracellular domain responsible for EGF binding. It was proposed that transformation of cells by AEV was the result, at least in part, of expression of a truncated EGF receptor molecule lacking the extracellular control domain[14].

We have now investigated EGF receptor expression in placental and A431 carcinoma cells[16]. A431 cells were chosen because they show two unusual features. First, they possess about 10–50 times more EGF receptors on their surface than most other cell types[16,17]. These high EGF receptor levels seem to be associated with the presence of a chromosomal translocation (M4) involving chromosome seven[18]. Second, their mitogenic response to EGF does not correlate with their increased EGF binding capacity. Thus while low concentrations of EGF stimulate their proliferation, EGF concentrations mitogenic for many cell lines inhibit proliferation of A431 cells[19,20]. It was therefore of great interest to investigate the genetic basis for the unusual features of this tumour cell line.

The complete amino acid sequence derived from human EGF receptor cDNAs from placental and A431 cells confirms the findings of Downward *et al.*[14], and matches partial amino acid sequences obtained from purified placental and A431 cell EGF receptor protein. Our findings also demonstrate that the EGF receptor gene is amplified in A431 cells. Furthermore, we describe the detection of an aberrant A431-specific chimaeric mRNA molecule which consists of a fragment of EGF receptor coding sequences recombined with sequences of unknown origin and function. This aberrant mRNA encodes only the external EGF binding domain of the receptor and lacks sequences which would encode the transmembrane and cytoplasmic domains. Expression of a truncated domain similar to the erb-B polypeptide found in AEV transformed cells[14] was not detected but its existence at low concentrations cannot be excluded.

Partial amino acid sequence

Receptor purified by immunoaffinity chromatography followed by gel permeation HPLC in guanidine after reduction and alkylation to cleave disulphide bonds was digested with trypsin and the peptides fractionated by preparative reverse-phase HPLC[14]. The purification and partial sequence analysis data of these peptides are shown in Fig. 1. The amino-terminal sequence of the intact receptor was deduced for a short stretch of 8 residues (see Fig. 1*b*). These partial amino acid sequences were found to be identical (apart from 2 residues) to those subsequently established by nucleotide sequencing of cDNA.

EGF receptor cDNA clones

Oligo(dT)-primed cDNA libraries were prepared from mRNAs derived from human placental tissue and epidermoid carcinoma (A431) cells using the λgt10 vector system[21]. As analysis of EGF receptor biosynthesis in A431 cells had suggested a polypeptide

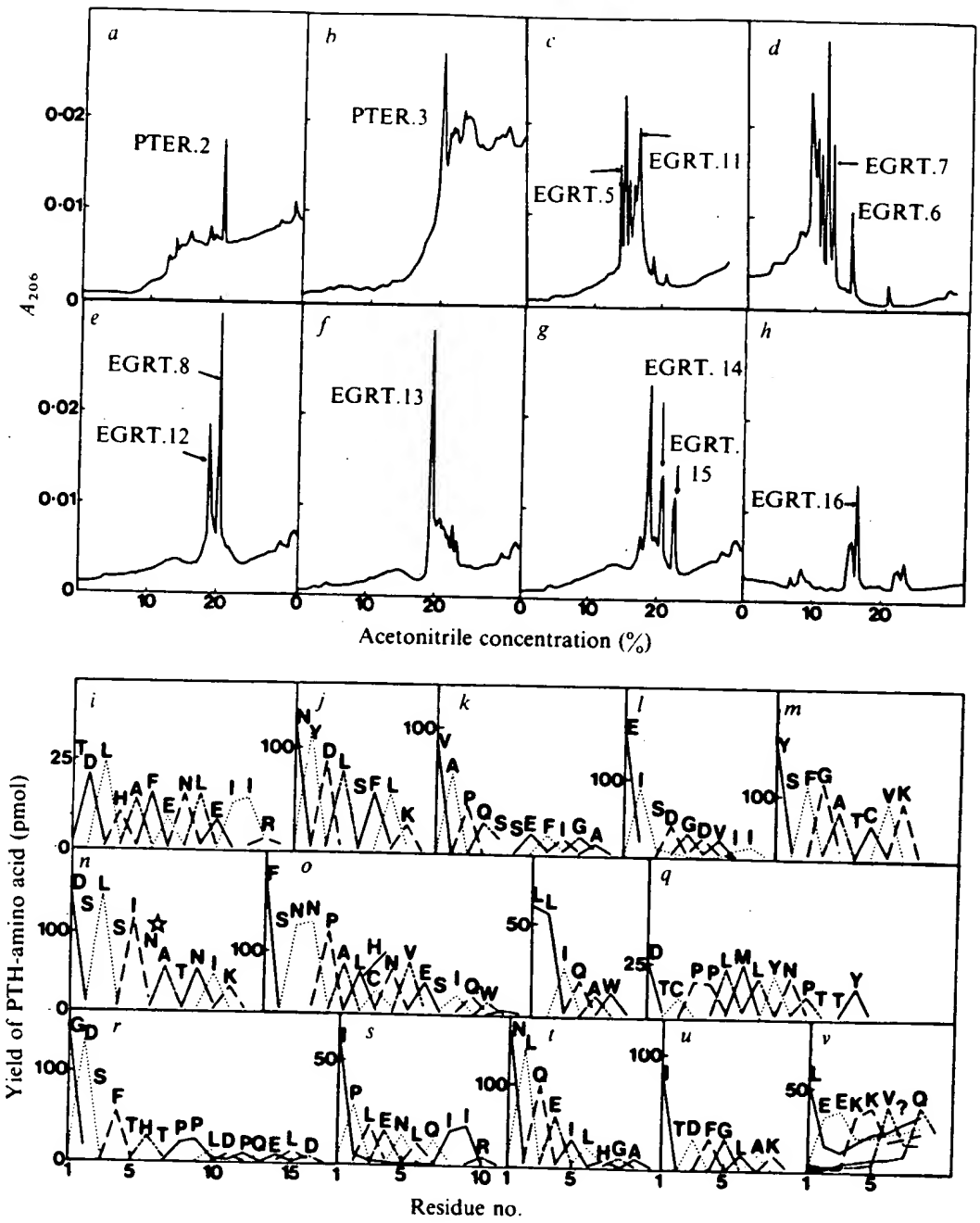
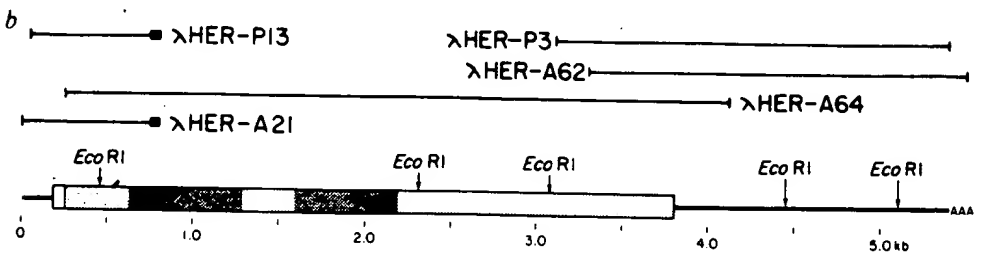


Fig. 1 EGF receptor peptide sequence analysis. *a–h*, Purification of peptides from the EGF receptor for sequence analysis. EGF receptor was purified from human placenta (*a*, *b*) or A431 human epidermoid carcinoma cells (*c–h*) using monoclonal antibody R1 (ref. 42); it was then fully reduced and alkylated, run on a gel permeation column, digested with trypsin and the resulting peptides fractionated by reverse-phase HPLC at *p*H 2.0, as described previously[14]. Fractions corresponding to optical density peaks in the first reverse-phase separation were further purified by reverse-phase HPLC on a Synchropak RPP C_{18} column (4.6 × 75 mm; Synchrom, Linden, Indiana) equilibrated in 10 mM ammonium acetate buffer, *p*H 6.5. A linear gradient of 0–45% acetonitrile over 45 min was used to elute peptides, flow rate 1 ml min^{-1}. The absorbance of the eluate was monitored at 206 nm and 0.5 ml fractions were collected. The figure shows the resulting chromatograms when the following fractions from the *p*H 2.0 fractionation were run: *a*, tryptic peptides from the placental EGF receptor eluting at 19% acetonitrile; *b*, placental peptides eluting at 22% acetonitrile; *c*, tryptic peptides from the A431 cell EGF receptor eluting at 21% acetonitrile; *d–h*, A431 peptides eluting at 18, 24, 26, 28 and 22% acetonitrile, respectively. *i–v*, Sequence analysis of peptides from the EGF receptor. Sequence determination was carried out as described[14]. The figure shows the yield of PTH derivatives of amino acids at each cycle of the Edman degradation for the peptides: *i*, PTER.2; *j*, PTER.3; *k*, EGRT.3 (for which the full purification has been described elsewhere[14]); *l*, EGRT.5; *m*, EGRT.6; *n*, EGRT.7; *o*, EGRT.8; *p*, EGRT.11; *q*, EGRT.12; *r*, EGRT.13; *s*, EGRT.14; *t*, EGRT.15; *u*, EGRT.16. *v*, The amino-terminal sequence for the whole EGF receptor molecule from A431 cells. The protein was treated with 0.05% SDS before sequencing. The seventh residue could not be assigned. ☆, The sixth residue of EGRT.7 was assigned from amino acid analysis of the peptide; the absence of asparagine-PTH at this sequencer cycle suggests that the residue is glycosylated.

backbone of molecular weight (MW) ~ 138,000 (ref. 22), at least 3,500 base pairs (bp) of coding sequence would be required, thus a sizing step was introduced to enrich for cDNA molecules greater than 2,000 bp long. For screening purposes a 51 bp-long DNA hybridization probe (Fig. 2*a*) was designed and synthesized[23] on the basis of a preliminary partial amino acid sequence of an A431 EGF receptor cyanogen bromide (CNBr) peptide (EGRC.1)[14]. Previous studies have shown that the use of a single long probe for clone screening under conditions of low stringency is often superior to the more commonly used technique which involves screening with pools of short oligonucleotide probes[24–26].

Initial screening of the placental cDNA library (2×10^5 clones) yielded seven distinct positive clones as judged by *Eco*RI restriction endonuclease mapping and Southern hybridization[27] with our synthetic probe. The longest fragment (1.4 kilobases, kb) of clone λ HER-P3 was sequenced using the dideoxynucleotide chain termination procedure[28–30]. The resulting sequence confirmed the presence of a region of similarity with our probe which, despite two errors in the preliminary determination of the amino terminus of the CNBr peptide, differed at only five positions from the sequence of the cloned cDNA (Fig. 2*a*, *c*). A second sequence analysis of this peptide gave the amino-terminal sequence Gly-Asp- (see ref. 14) and reinspection of these data revealed a trace of PTH-Asp at step 1 supporting the sequence found in analysis of cDNA clones. Comparison of the



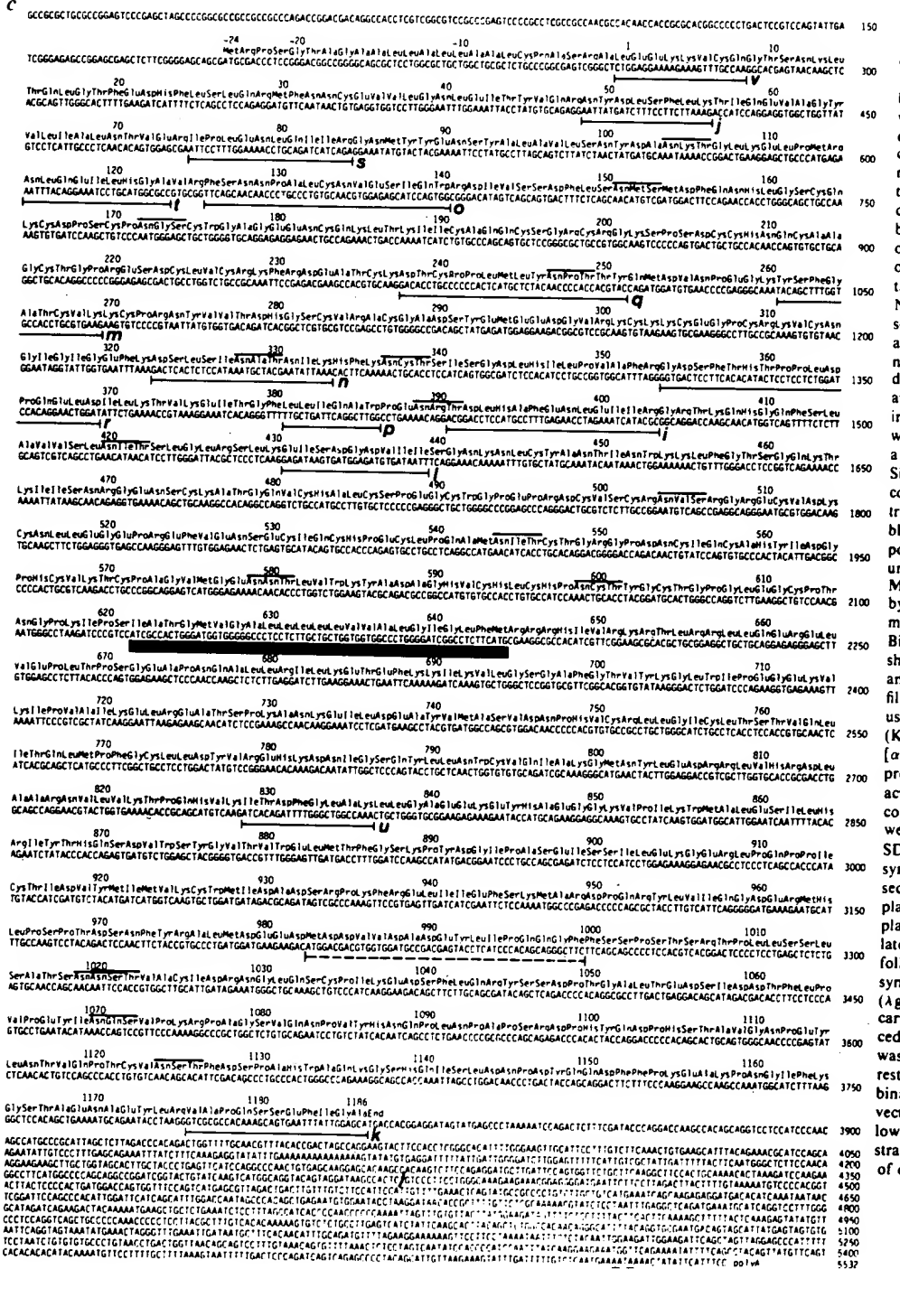


Fig. 2 EGF receptor cDNA nucleotide sequence and predicted amino acid sequence. *a*, The nucleotide sequence of probes based on the amino acid sequence of a CNBr peptide of A431 EGF receptor. The initial amino acid residue was assumed to be methionine since the peptide was generated by cyanogen bromide cleavage[18]. The DNA sequence was designed on the basis of published codon usage frequencies[57]. Two overlapping 30 residue-long oligonucleotides (underlined) from coding and non-coding strands were chemically synthesized[23]. *b*, Schematic representation of human EGF receptor cDNA clones. Overlapping clones from placenta ("P") and A431 ("A") epidermoid carcinoma cells used in sequence determination and a schematic diagram of the complete cDNA structure. Untranslated sequences are represented by a line, coding sequences are boxed. The white portion represents sequences which encode the signal peptide; hatched regions code for the mature receptor, dark hatched regions are high in cysteine residues (see text). The black region demarcates the coding region for the putative transmembrane domain. Black boxes on the 3' ends of λ clones indicate that these clones were obtained by specific priming (complementary sequence at positions 757–780). *c*, Nucleotide and predicted amino acid sequence of EGF receptor. Nucleotides are numbered at the right; amino acids are numbered throughout. Peptide sequences derived from purified receptor preparations are underlined and labelled according to Fig. 1(*i–u*). The peptide underlined with a broken line was that used to design a long synthetic DNA probe (see Fig. 2*a*). Sites of potential asparagine-linked glycosylation are overlined; the putative transmembrane region is demarcated by a black bar. The AATAAA box close to the polyadenylated 3' end of the mRNA is underlined.

Methods: *a*, The probe was radiolabelled by first phosphorylating the synthetic 30-mers using T4 polynucleotide kinase (PL Biochemicals) and [γ-^{32}P]ATP (Amersham, >5,000 Ci mmol^{-1}) followed by annealing the kinased single strands and filling in the complementary sequences using *Escherichia coli* DNA polymerase I (Klenow fragment) with [α-^{32}P]dCTP and [α-^{32}P]dATP (Amersham) yielding probes of 5×10^8 c.p.m. per μg specific activity. Hybridization was carried out in conditions published elsewhere[24]. Filters were washed at 45 °C with 0.2×SSC, 0.1% SDS. Asterisks denote nucleotides of the synthetic probe which did not match the sequence of the isolated clone. *b*, Cytoplasmic mRNA from total human term placental tissue and A431 cells were isolated in the presence of RNase inhibitors following standard protocols[58]. cDNA synthesis[59], cloning into a λ vector (λgt10)[21], and plaque screening[60] were carried out according to published procedures. *c*, Nucleotide sequence analysis was carried out by subcloning suitable restriction fragments from our recombinant λ clones into M13-based cloning vector (M13mp10 and M13mp11[29]) followed by primed DNA synthesis on single-stranded DNA templates in the presence of dideoxynucleoside triphosphates[28,30].

```
HER      596  PECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVKTCPAGVMGENNTLVWKYADAGHVCHL
v-erb-B                        MK  F       A   L  D  R     NA  Q

HER      596  CHPNCTYGCTGPGLEGCPTNGPKIPSIATGMVGALLLLLVVALGIGLFMRRRHIVRKRTL
v-erb-B           R  K        -  S  T   A  V  G  C  V  G      YL

HER      656  RRLLQERELVEPLTPSGEAPNQALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVK
v-erb-B                              H          V          I

HER      716  IPVAIKELREATSPKANKEILDEAYVMASVDNPHVCRLLGICLTSTVQLITQLMPFGCLL
v-erb-B                                                          Y

HER      776  DYVREHKDNIGSQYLLNWCVQIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLA
v-erb-B           I                          E

HER      836  KLLGAEEKEYHAEGGKVPIKWMALESILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIP
v-erb-B            D

HER      896  ASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKFRELIIEFSKMARDPQRYL
v-erb-B            V                                        A        P

HER      956  VIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIPQQGFFSSPSTSRTPLLSSLS
v-erb-B                       K  T  E   E  I       V  H    N

HER     1016  ATSNNSTVACIDRNGLQSCPIKEDSFLQRYSSDPTGALTEDSIDDTFLPVPEYINQSVPK
v-erb-B             ATN       -  GH  VR    V          NFL  E    G   A   V  LM

HER     1076  RPAG-SVQNPVYHNQPLNPAPSR---QPHYQDPHSTAVGNPEYLNTVQPTCVNSTFDSPA
v-erb-B       K  STAM   QI  -  FISLT  I  KLPM  SR   NS      D        N  SPLAKTV  E  SP

HER     1132  HWAQKGSHQISLDNPDYQQDFFPKEAKPNGIFKGSTAENAEYLRVAPQSSEFIGA
v-erb-B       Y  I  S  N    N             L  TSCS
```

Fig. 3 Comparison of EGF receptor with v-*erb-B* predicted amino acid sequences. Only non-matching amino acids are shown for v-*erb-B*. The bracket on the first line indicates the start of the v-*erb-B* sequence[15]. The putative transmembrane region is boxed; the arrow demarcates the autophosphorylated tyrosine residue in pp60^{src} (ref. 61). Asterisks indicate residues conserved between the receptor sequence and several members of the *src* gene family (*src*, *fes*, *yes*, *fps*). Hyphens (gaps) have been introduced for optimal alignment. The C-terminal tetrapeptide of the v-*erb-B* sequence derived from the AEV *env* gene is underlined[15].

complete sequence of the 2.3-kb cDNA clone λHER-P3 with the nucleotide sequence of the oncogene v-*erb-B*[15] showed that it coded for 243 carboxy-terminal amino acids of the receptor followed by a termination codon and about 1,600 bp of presumably 3′-untranslated sequences lacking a poly(A) homopolymer stretch (see Fig. 2*b*).

Using the 1.4 kb *Eco*RI fragment of λHER-P3 as a probe, 21 clones were isolated from an A431 cDNA library and characterized by *Eco*RI restriction analysis. Clones λHER-A62 and λHER-A64 were found to contain overlapping DNA fragments, which together spanned a stretch of 5.2 kb (Fig. 2*b*) of DNA. Complete sequence analysis revealed an open reading frame coding for 1,186 amino acid residues, including carboxy-terminal sequences identical to λHER-P3 and very similar to v-*erb-B* transforming protein. Clone λHER-A62 was found to contain the complete 3′-untranslated region (1,913 bp), as evidenced by the presence of an AATAAA sequence[31] followed by a poly(A) tail 15 nucleotides downstream.

The amino acid sequences of 13 tryptic peptides as well as those of the 6 previously reported[14] were located within the predicted amino acid sequence (see Figs 1, 2*c*). Only two differences were detected between the 230 amino acid residues determined by peptide sequencing and those derived from analysis of the cloned cDNA (Cys as opposed to His at position 133). Reinspection of amino acid sequencing data revealed a failure to distinguish between Cys and His in peptide EGRT.8 and to assign amino-terminal Asp in peptide EGRC.1. Remarkably, clone λHER-A64 starts with a nucleotide sequence encoding the amino acid sequence NH_2-Leu-Glu-Gly-Lys-Lys, which is identical to the amino-terminal sequence of the A431 receptor polypeptide (see Figs 1*c*, 2*c*).

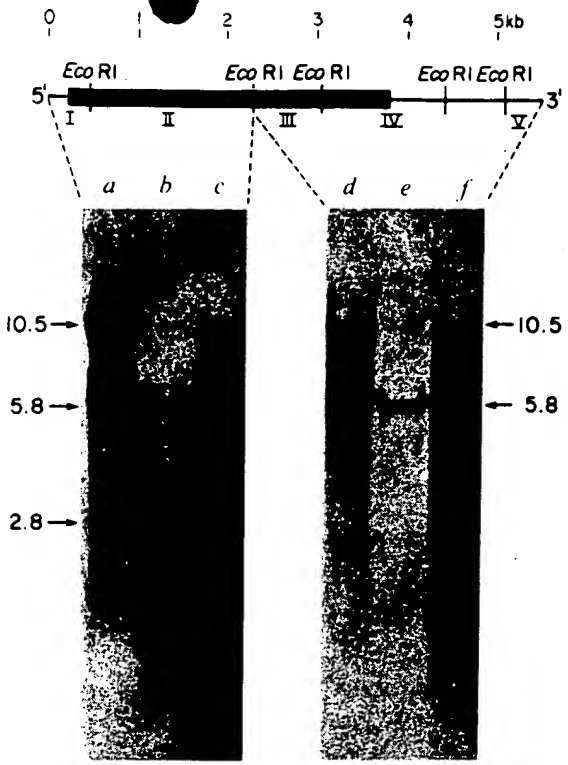


Fig. 4 Northern blot analysis of cytoplasmic placental and A431 mRNAs. Lanes *a* and *d*, placental mRNA; *b* and *e*, A431 mRNA; *c* and *f*, mRNA from a second, independently passaged line of A431 cells. Dashed lines indicate the regions of the schematically depicted cDNA giving rise to the hybridization patterns shown. Numbers on the left and right indicate sizes (in kb) of major mRNA bands. Roman numerals denote *Eco*RI fragments used as hybridization probes (see below).

Methods: A431 cells were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum in an atmosphere of 10% CO_2–90% air at 37 °C. Cytoplasmic mRNA isolated from frozen placental tissue after pulverization in liquid N_2 and from fresh A431 cells[58] was separated (2 μg per lane) on 1% agarose formaldehyde gels and transferred to nitrocellulose filters as described by Lehrach *et al.*[36]. Rat liver ribosomal RNA (18S = 1,874 nucleotides; 28S = 4,718 nucleotides; I. Wool, personal communication) and 42S vesicular stomatitis virus RNA (~12 kb) were used as size markers. Nitrocellulose filters were hybridized under high-stringency conditions with radiolabelled *Eco*RI fragments (I = 293 bp; II = 1,838 bp; III = 768 bp; IV = 1,373 bp; V = 432 bp) derived from clones λHER-A62 and λHER-A64.

In order to obtain missing sequences upstream from those which encode the amino-terminus of the mature protein, an A431 cDNA library generated by priming with a 24 nucleotide-long synthetic oligomer (see Fig. 2 legend) was screened with the 5′-most *Eco*RI fragment of λHER-A64. Clone λHER-A21 perfectly matched the sequence of λHER-A64 and extended it 260 bp further upstream. The new sequence contained coding information for an additional 24 predominantly hydrophobic amino acids, including the presumptive initiator methionine. This sequence probably represents the EGF receptor signal peptide which is necessary for transfer of the nascent polypeptide into the lumen of the endoplasmic reticulum. Basic and polar amino acid residues flank the hydrophobic core of the amino-terminal 24-amino acid sequence, a common structural feature of signal peptides. Furthermore, within the 186 bp of 5′-untranslated sequence, an in-frame termination codon is located 36 nucleotides upstream (Fig. 2*c*), supporting the assignment of residue −24 as the initiation methionine of the 1,210 amino acid-long EGF receptor precursor polypeptide. Placental EGF receptor clones were generated in an analogous priming experiment. Clone λHER-P13 (Fig. 2*b*) was characterized and matched perfectly the sequence of λHER-A21. These data show

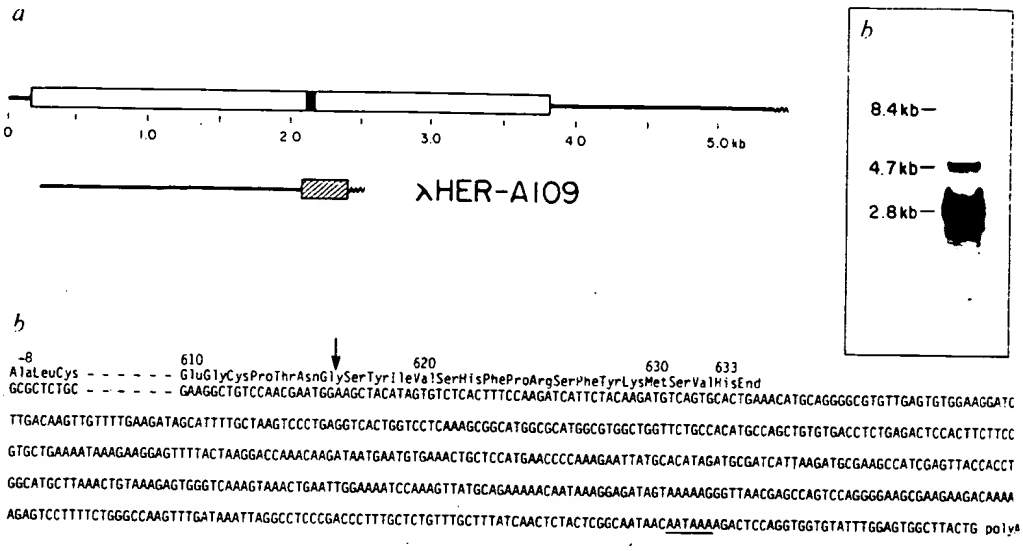


Fig. 5 Structure of an A431 cell-specific cDNA variant. *a*, Schematic map of normal (top) and variant (bottom) cDNA. The solid box represents transmembrane sequences. The black line in the variant cDNA clone λHER-A109 represents the region of identity with the normal EGF receptor sequence; the cross-hatched box represents a region of unknown origin. ~~, Poly (A) tails. *b*, Northern blot analysis using a probe derived from the non-homologous region of variant cDNA clone λHER-A109. A 260-bp *Hpa*I–*Eco*RI fragment was used as radioactive probe (10^8 c.p.m. μg^{-1}). Blots were exposed for 72 h. Sizes of mRNAs are shown in kb. 28S (4,718 nucleotides) and 18S (1,874 nucleotides) ribosomal RNAs were use as size standards. No hybridization was detected in a parallel experiment with placental mRNA. *c*, Nucleotide sequence of cloned cDNA from λHER-A109. The start of the cDNA is shown, together with the junction of non-homology to EGF receptor cDNA and the entire non-homologous 3′ end.

that both placental and A431 receptors are not synthesized as larger precursor molecules aside from the signal peptide; the unmodified precursor and mature polypeptides have predicted molecular weights of 134,300 and 131,360, respectively. These values are in good agreement with the molecular weight (138,000) determined for EGF receptor synthesized in tunicamycin-treated A431 cells, where *N*-linked glycosylation does not occur[22].

Human EGF receptor/v-*erb-B* homology

Partial amino acid sequence analysis of peptides obtained from placental and A431 cell EGF receptor preparations had demonstrated that the EGF receptor matches, in part, the amino acid sequence of the product of the v-*erb-B* oncogene from avian erythroblastosis virus (AEV-H[14,15]). Availability of the complete EGF receptor sequence derived from cloned cDNA now permits a detailed sequence comparison. Figure 3 shows an alignment of the 604 residue-long v-*erb-B* sequence[15] with 631 carboxy-terminal amino acids of the human EGF receptor. Extensive amino acid sequence homology is detected over a 376 residue-long core region, beginning at the cytoplasmic junction with the transmembrane domain present in both sequences; 95% of the residues in this region are conserved, which is highly significant considering the evolutionary distance between birds and humans, and therefore probably represent sequences essential for the function of this domain. A 244 residue-long stretch within this region contains 60 residues (24.6%) which are conserved in other oncogenes of the *src* gene family (for clarity, only *src*, *fes*, *yes* and *fps* are shown in Fig. 3)[32]. These include the tyrosine which is the substrate for the autophosphorylation reaction of pp60^{src} (refs 33, 34; Fig. 3). The equivalent region of the EGF receptor also retains the amino acid residues thought to be involved in the formation of an ATP binding site[32,35]. The extent of the sequence conservation is reduced in sequences flanking this 376-amino acid core region until, at residue 600, the v-*erb-B* sequence is fused to the carboxy-terminal tetrapeptide of the AEV *env* protein[15]. The EGF receptor sequence continues from the point of divergence at the 3′ end for another 32 amino acids. In the light of the sequence identity in the putative tyrosine specific kinase domain, it is surprising that the *erb-B* transforming protein has not been shown to possess kinase activity. This may reflect technical difficulties or could indicate that the carboxy-terminal region, in which there is less similarity between the sequences, may be important for expression of the EGF receptor kinase activity. The complete sequence further substantiates the hypothesis that v-*erb-B* oncogene sequences originated from cellular sequences which encode the avian EGF receptor[14]. As the nucleotide and hence the amino acid sequence of the avian EGF receptor is unknown, the relevance of particular amino acid differences between the human receptor and avian oncogene sequences to the transforming function of the virus remains unclear. The extent of similarity further supports our previous suggestion[14] that other oncogene products that have regions of amino acid sequence similar to v-*erb-B* and the EGF receptor, which form a group known as the *src* gene family, may be derived from cellular sequences encoding other receptors having tyrosine kinase activities such as the PDGF, IGF-I and insulin receptors.

Receptor-related mRNAs

To determine the size of the mRNA encoding the EGF receptor, Northern blot hybridization[36] experiments were done using cytoplasmic mRNA isolated from term placenta and from two independently propagated lines of A431 cells. The RNA preparations were probed with various segments of the cloned EGF receptor cDNA. Figure 4 shows that different patterns of hybridization were obtained with probes derived from sequences coding for the 5′-untranslated region and the extracellular domain as compared with probes derived from the cytoplasmic domain and the 3′-untranslated regions. In all experiments, two transcripts of 5.8 and 10.5 kb were detected in both A431 and placental RNA preparations, however the relative amounts of the transcripts differed between placenta and A431 cells, as well as between the two lines of A431 cells. Similarly, two transcripts of 12 and 9 kb have been detected in normal chicken embryo mRNA using v-*erb-B* probes[37]. These two transcripts could result from transcription of two closely related genes, or, alternatively, represent transcripts of different length from the same gene. Transcription of the mouse α-amylase and dihydrofolate reductase genes, for example, involves usage of multiple poly-(A) addition sites, generating mRNAs of variable lengths from a single gene[38,39]. Alternatively, variable splicing events could generate two mRNAs, perhaps with different function, from a single gene[40,41]. The variation in the relative levels of the two transcripts in placenta and the two A431 cell lines suggests that some specific mechanism may regulate their relative abundance depending on cell type, physiological condition or developmental state. Further experiments are needed to elucidate this phenomenon.

Probes derived from the 5′ half of the cDNA hybridize to a 2.8-kb mRNA present only in A431 cells. Compared with the

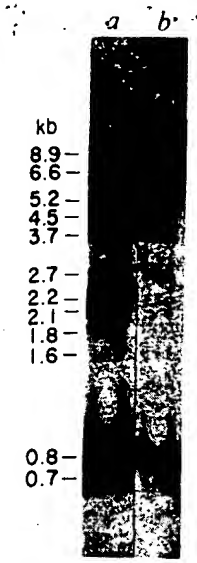


Fig. 6 Southern blot analysis of A431 (*a*) versus placental (*b*) DNA. High-molecular weight chromosomal DNA was isolated as described[58]. DNA (10 μg) was digested with excess *Eco*RI (New England Biolabs) then analysed by Southern blot hybridization. A mixture of all *Eco*RI fragments of λHER-A62 and λHER-A64 (as in Fig. 4) was radiolabelled with [α-^{32}P]dATP and [α-^{32}P]dCTP (Amersham) by the procedure of Taylor *et al.*[62]. Processed filters were exposed to X-ray film for 5 days at −70 °C using Cronex lightning plus intensifier screens. λ phage DNA was digested with *Eco*RI and *Hin*dIII was used as a size marker.

two normal mRNAs, this RNA species is overexpressed ~100-fold (Fig. 4*b*, *c*). In addition to the normal mRNA bands at 10.5 and 5.8 kb, various fainter bands are observed after long exposure times in both placenta and A431 RNA preparations (see Fig. 4 and data not shown). A 6.1-kb mRNA which is present in both A431 cells and placenta appears to be homologous to all probes used; the same is true for a placenta-specific 13-kb mRNA. Another group of mRNAs from both sources hybridizes with either 5′ (1.8 kb) or 3′ (8.9 kb) probes. Faint A431-specific hybridization signals are detected with 5′ probes at 3.3, 1.65, 1.5 and 1.3 kb in addition to the predominant 2.8-kb mRNA. These bands may represent low-abundance variant transcripts of the same gene, or possibly transcripts of other genes which are structurally related to either the entire EGF receptor or one of its major domains.

Truncated EGF receptor sequences

To characterize in detail the predominant 2.8-kb variant mRNA from A431 cells, an oligo(dT)-primed cDNA bank (>2 kb; ~10^5 clones) was screened with a 5′-specific EGF receptor cDNA probe (Fig. 4, fragment 1). Of 167 positives, the three clones with the largest inserts were found by restriction enzyme analysis to be related to but distinct from clone λHER-A64. Clone λHER-A109 (2.5 kb) was sequenced and was found to be identical with our EGF receptor cDNA sequence starting at amino acid position −8 in the signal sequence and extending through almost the entire extracellular domain of the receptor. The amino acid sequence identity terminated abruptly, however, at amino acid residue 615, shortly before the transmembrane portion of the normal sequence (see Fig. 5*a*, *c*). Downstream from this residue the open reading frame continues for another 18 amino acids, followed by an opal codon (TGA) and a 548-nucleotide 3′-untranslated sequence which includes an AATAAA box 34 nucleotides upstream from the poly(A) addition site. The cDNA (λHER-A109) encodes an approximately 70,000-MW truncated EGF receptor protein which, because it contains a signal peptide but lacks the receptor transmembrane region, should be secreted by A431 cells. Monoclonal antibody R1 (ref. 42), which recognizes the extracellular domain of the EGF receptor, does in fact immunoprecipitate a secreted 115,000-MW glycopeptide, which is thought to be formed by post-translational modification of a 68,000-MW polypeptide[22]. Analogous truncation of the vesicular stomatis virus transmembrane glycoprotein by molecular cloning techniques led to secretion of a soluble protein, albeit at a severely reduced rate[43].

To establish the identity of this chimaeric cDNA, differential Northern blot hybridization experiments were carried out. Hybridization using a 185-bp *Eco*RI–*Sac*I probe derived from clone λ HER-A21, which includes only sequences upstream from the 5′ end of the variant clone λ HER-A109, resulted in a pattern identical to that seen in Fig. 4*b*, *c*. This finding, in combination with the size and sequence of clone λ HER-A109, showed that the 2.8-kb mRNA shares 5′ sequences with the 5.8- and 10.5-kb mRNAs. It is not yet clear whether transcription of these mRNAs is under the control of the same regulatory elements or whether sequence rearrangements in the 5′-flanking region of the receptor gene lead to overexpression of the 2.8-kb mRNA. A variant-specific probe (260 bp *Hpa*I–*Eco*I) derived from the 3′ end of clone λ HER-A109 hybridized to the major A431-specific 2.8-kb band as well as to fainter 4.7- and 8.4-kb bands, but not to the 5.8- and 10.5-kb mRNAs (Fig. 5*b*). Although the two less abundant (4.7- and 8.4-kb) A431-specific transcripts have the same 3′ untranslated sequence as our chimaeric cDNA clone λ HER-A109 and therefore could have engendered the cDNA, since neither of them showed homology with various EGF receptor gene sequences, we conclude that our chimaeric cDNA was derived from the 2.8-kb mRNA and that the 4.7- and 8.4-kb mRNAs probably represent transcripts from the parental gene which contain the 3′ region of the 2.8-kb mRNA.

Gene amplification

A431 epidermoid carcinoma cells express about 10–50 times more EGF receptors on their surfaces than most other cell types[16,17]. The remarkably high EGF binding capacity of these cells seems to be correlated with the presence of chromosomal translocations involving chromosome 7 (ref. 18), which is known to carry the EGF receptor gene[44,45]. In addition to the two translocated fragments of chromosome 7 (M4 and M14), A431 cells contain two copies of the normal chromosome 7 (ref. 18). To investigate further the genetic basis for the high EGF receptor levels in these cells, we compared A431 cell chromosomal DNA with DNA from placenta and other normal tissues. Assuming that the A431 cell chromosomal translocations described by Shimizu and Kondo[18] involve the entire EGF receptor gene, the signal in Southern hybridization experiments for A431 DNA was expected to be twice that detected for DNA from placenta and other human tissues. Figure 6 shows an autoradiograph of a Southern blot hybridization analysis of equal amounts of *Eco*RI-digested A431 (lane *a*) and placenta (lane *b*) DNAs. On the basis of densitometric quantification we estimate that the EGF receptor gene is amplified 15–20 times in A431 cells compared with DNA from normal diploid cells. Variable intensities of hybridization signals were observed for different bands within one lane, which may be due to different factors such as efficiency of transfer and binding to nitrocellulose, co-migration of hybridizing DNA fragments and the use of a mixed probe consisting of five DNA fragments of different size (see Fig. 6 legend). The highly selective conditions used in the hybridization experiments excluded cross-hybridization to related but non-identical sequences. Based on more detailed Southern hybridization experiments using various regions of the cDNA to probe separate blots, we estimate that the gene coding for the 5.8-kb EGF receptor mature transcript is greater than 50 kb in size (data not shown).

Our data suggest that various segments of the entire gene are amplified to the same extent. Furthermore it appears that the 8.9-kb hybridization signal obtained with A431 DNA (Fig. 6, lane *a*) is due to sequence rearrangements, since it is not present in placental and other normal DNAs (Fig. 6 and unpublished results). Additional indications of sequence rearrangements affecting the amplified receptor gene and its flanking regions have been detected in genomic Southern hybridization experiments using the 3′ nonhomologous sequence of clone λHER-A109 as a probe (A.U., in preparation).

Discussion

We have presented here for the first time the complete amino acid sequence of a cell surface receptor for a peptide ligand—the mitogen EGF.

Inspection of the receptor's predicted primary sequence reveals a stretch of 23 predominantly hydrophobic amino acids

(residues 622–644) which contains only a single amino acid with a side chain capable of hydrogen bonding. This sequence, in an α-helical configuration, could span the lipid bilayer and the group of basic residues immediately C-terminal (residues 89–101) are likely to interact with the phospholipid head groups in the membrane. It is unlikely that the external domain of the receptor contains regions that cross the lipid bilayer since in A431 cells a truncated receptor resulting from gene rearrangement and consisting of almost the entire external domain is secreted. The cytoplasmic domain may interact with the lipid bilayer but at regions with different structural features from those of the 23-amino acid transmembrane domain. The receptor therefore can be divided into two functional domains—an extracellular one of 621 amino acids and a cytoplasmic domain of 542 amino acids—linked by a short transmembrane region.

The extracellular domain would contain the EGF-binding site. This domain contains 12 of the 16 possible sites for asparagine-linked glycosylation (Asn-X-Ser or Asn-X-Thr). It is likely that the majority of these sites are indeed modified since Mayes and Waterfield[22] have shown that 11 sites are glycosylated in the precursor of the secreted external domain of the molecule.

A particularly striking feature of the receptor concerns the distribution of cysteine residues between the extracellular and cytoplasmic domains. The cytoplasmic domain of 542 amino acids contains nine cysteine residues (~2%), a value well within the range of most cytoplasmic proteins. The extracellular domain of 621 amino acids, however, contains 51 cysteine residues, most of which are concentrated within two regions of ~170 amino acids located at positions 134–313 and 446–612 (see Fig. 2*c*). Alignment of these cysteine-rich sequences with one another reveals similarities in the spacing of the individual cysteine residues, possibly reflecting the involvement of these residues in the formation of two repeated structures and a common evolutionary origin by gene duplication for this part of the receptor. We do not yet know whether these cysteine residues are all disulphide-bonded.

The cytoplasmic domain undoubtedly contains the tyrosine-specific protein kinase activity and has the stretch of amino acid sequence (approximately residues 690–940) that is shared by transforming proteins of the *src* family.

The question remains, how do the domains function in signal transduction across the membrane? If, as we suggest, the EGF receptor polypeptide crosses the membrane only once, then a signal must be transduced through this region unless two or more receptor molecules are involved or an additional polypeptide(s) mediates the response to EGF. What could be the nature of the signal? Since EGF induces rapid changes in ion movements[3] it is possible that the receptor itself could function as a channel. Very little is known about structure–function relationships for such channels except in the case of bacteriorhodopsin[46] and the acetylcholine receptor[47]. For both these proteins it seems that multiple transmembrane segments formed from either a single polypeptide in the case of bacteriorhodopsin or from multiple subunits in the case of the acetylcholine receptor contribute to channel function. Inspection of the putative EGF receptor transmembrane sequence suggests that, even if ligand-induced aggregation brought several transmembrane segments into close apposition, the sequence itself does not contain any amino acids with side chains that could be involved in ion or proton translocation.

An alternative hypothesis is that an EGF-induced conformational change in the external domain is transmitted to the cytoplasmic domain, perhaps as a result of movement of the transmembrane domain but more likely involving receptor aggregation. EGF will induce receptor clustering[48–50] and antibody cross-linking can mimic the effects of the growth factor[51,52]. Such clustering of extracellular domains would necessarily bring the corresponding cytoplasmic domains into close proximity, allowing inter-receptor interactions which might include stimulation of tyrosine kinase activity resulting in autophosphorylation. The experimental evidence cannot exclude a mechanism where a receptor molecule bearing bound ligand interacts with other polypeptides which transduce the transmembrane signal and hence stimulate kinase or other activities indirectly. However it is important to note that EGF can stimulate tyrosine kinase activity of highly purified EGF receptor (ref. 53 and Y.Y., in preparation).

Comparison of receptor gene expression patterns in normal placental tissue and the tumour cell line A431 has revealed distinct differences both at the transcriptional and at the gene structural level. We have shown that the EGF receptor is encoded by a 5.8-kb mRNA in both placental and A431 cells. We also detect a 10.5-kb mRNA which may be a transcriptional variant derived from the same gene, perhaps carrying a longer 3′-untranslated region, analogous to other genes[38,39]. Both the 5.8- and 10.5-kb mRNAs are expressed at variable but similar levels in both cells, which is remarkable because A431 cells are known to possess 10–50-fold more cell surface EGF-binding sites and associated tyrosine kinase activity than most other cell types[16,17]. The abnormality we detect instead is that A431 cells synthesize a 2.8-kb mRNA at levels 100-fold greater than that of the larger mRNA species. This mRNA encodes a 70,000-MW secreted polypeptide representing almost the entire extracellular EGF binding domain of the receptor[22]. Since this truncated receptor is secreted[22], it cannot account for the observed overexpression of EGF receptors detected by increased EGF binding capacity[1], monoclonal antibody binding[42], immunoprecipitation of tyrosine kinase activity[4,5] and receptor protein quantification as observed on SDS-polyacrylamide gel analysis of these cells[14].

Thus the important question of the mechanism for the dramatically increased EGF receptor levels in A431 cells remains unsolved. We have found that the EGF receptor gene is amplified 15–20-fold in A431 cells, yet amplification does not lead to concomitant increased synthesis of the 5.8- and 10.5-kb mRNAs. A431 cells contain two normal copies of chromosome seven and two chromosomes bearing a chromosome seven translocation[18]. Others have observed that gene amplification events are often associated with chromosomal translocations[54,55], and Shimizu and Kondo have shown that the overproduction of EGF receptor in A431 cells is genetically linked to the chromosome seven translocation (M4)[18]. We propose that the 5.8-kb and 10.5-kb mRNAs are derived from the two normal copies of chromosome seven, while the novel overexpressed 2.8-kb mRNA is derived from the amplified gene copies. Our results do not explain the link between high levels of intact receptors with chromosome 7 translocation.

Amplification of the EGF receptor gene in A431 cells could have occurred intrachromosomally or by generation of acentric extrachromosomal elements (so-called double minute chromosomes). It is possible that the sequence rearrangements we detected (A.U., in preparation) occurred during or after the translocation process and created an altered copy of the EGF receptor gene, leading to truncation and moderate overexpression, before amplification. Alternatively, sequence alterations could have occurred during gene amplification, resulting in overexpression of a single aberrant gene. Detailed analysis of the EGF receptor gene and flanking sequences from normal and A431 cells will be necessary to determine whether the occurrence of the amplified gene generates a truncated mRNA and why this mRNA is overexpressed.

It is also not clear from these results why A431 cells are uniquely unresponsive to the amounts of growth factor which are usually mitogenic for fibroblasts[19,20,56].

The finding that A431 vulval carcinoma cells produce a truncated polypeptide which is virtually identical to the external domain of the receptor is particularly intriguing in view of our earlier observation that the *v-erb-B* oncogene encodes what seems to be a truncated avian receptor polypeptide corresponding to the transmembrane and cytoplasmic domains[14]. We cannot yet exclude the presence, at a low level, in A431 cells of an mRNA which could encode such a cytoplasmic membrane-associated domain. It seems likely that the A431 2.8-kb mRNA and the secreted EGF receptor external domain are unique to

A431 cells and do not provide any clues to the neoplastic origin or character of this cell line. Since the A431 cell line—and the many variants of this cell line which must now exist (see Fig. 5)—has a multitude of defects exemplified by its 78 chromosomes[18], it is clearly best to regard this cell line with caution when exploring the mechanism of action of EGF or in characterizing changes which relate in a meaningful way to neoplasia.

The structural studies described here provide a basis from which to probe other cell lines and primary tumour tissue for alterations in the EGF receptor, its gene, or its transcription pattern.

After this manuscript was completed two publications related to the subject appeared. It was shown that A431 cells produce a 105,000-MW protein related to the external domain of the molecule[63], as previously reported by Mayes and Waterfield[22], and that amplification and altered expression of the EGF receptor gene occurs in A431 cells[64].

We thank Mark Vasser, Parkash Jhurani and Kim Dinkelspiel for synthesis of oligonucleotide primers; Jeanne Arch for manuscript preparation; and Drs Suzanne Pfeffer and S. Earp for stimulating discussions. This work was supported in part by grants from the NIH (CA 25820 to J.S.) and from the US–Israel Binational Foundation.

Received 24 April; accepted 10 May 1984.

1. Carpenter, G. *Handbk exp. Pharmac.* **57**, 89–132 (1981).
2. Guroff, G. (ed.) *Growth and Maturation Factors* Vol. 1 (Wiley, New York, 1983).
3. Rozengurt, E. *Molec. Biol. Med.* **1**, 169–181 (1983).
4. Cohen, S., Carpenter, G. & King, L. Jr *J. biol. Chem.* **255**, 4834–4842 (1980).
5. Ushiro, H. & Cohen, S. *J. biol. Chem.* **255**, 8363–8365 (1980).
6. Carpenter, G. & Cohen, S. *J. Cell Biol.* **71**, 154–171 (1976).
7. Schlessinger, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 2659–2663 (1978).
8. Pike, L. J. *et al. J. biol. Chem.* **258**, 9282–9390 (1983).
9. Heldin, C.-H., Ek, B. & Ronnstrand, L. *J. biol. Chem.* **258**, 10054–10061 (1983).
10. Kasuga, M. *et al. Nature* **298**, 667–669 (1982).
11. Roth, R. A. & Cassell, D. J. *Science* **219**, 299–301 (1983).
12. Jacobs, S. *et al. J. biol. Chem.* **258**, 9581–9584 (1983).
13. Bishop, J. M. *A. Rev. Biochem.* **52**, 301–354 (1983).
14. Downward, J. *et al. Nature* **307**, 521–527 (1984).
15. Yamamoto, T. *et al. Cell* **35**, 71–78 (1983).
16. Fabricant, R. N., De Larco, J. E. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 565–569 (1977).
17. Wrann, M. M. & Fox, C. F. *J. biol. Chem.* **254**, 8083–8086 (1979).
18. Shimizu, N. & Kondo, I. *Cytogenet. Cell Genet.* **32**, 316 (1982).
19. Gill, G. N. & Lazar, C. S. *Nature* **293**, 305–307 (1981).
20. Kawamoto, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 1337–1341 (1983).
21. Huynh, T., Young, R. & Davis, R. in *Practical Approaches in Biochemistry* (ed. Grover, D.) (IRL, Oxford, 1984).
22. Mayes, E. L. V. & Waterfield, M. D. *EMBO J.* **3**, 531–537 (1984).
23. Crea, R. & Horn, T. *Nucleic Acids Res.* **8**, 2331–2348 (1980).
24. Ullrich, A., Berman, C. H., Dull, T. J., Gray, A. & Lee, J. M. *EMBO J.* **3**, 361–364 (1984).
25. Anderson, S. & Kingston, I. B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6836–6842 (1983).
26. Jaye, M. *et al. Nucleic Acids Res.* **11**, 2325–2335 (1983).
27. Southern, E. *J. molec. Biol.* **98**, 503–517 (1975).
28. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
29. Messing, J. & Vieira, J. *Gene* **19**, 269–276 (1982).
30. Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 309–321 (1981).
31. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211–214 (1976).
32. Privalsky, M. L., Ralston, R. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 704–707 (1984).
33. Levinson, A. D. Oppermann, H., Varmus, H. E. & Bishop, J. M. *J. biol. Chem.* **255**, 11973–11980 (1980).
34. Erikson, R. L., Collett, M. S., Erikson, E. & Purchio, A. F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6260–6264 (1979).
35. Barker, W. C. & Dayhoff, M. D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2836–2839 (1982).
36. Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. *Biochemistry* **16**, 4743–4751 (1977).
37. Vennstrom, B. & Bishop, J. M. *Cell* **28**, 135–143 (1982).
38. Tosi, M., Young, R. A., Hagenbuchle, O. & Schibler, U. *Nucleic Acids Res.* **9**, 2313–2323 (1981).
39. Setzer, D. R., McGrogan, M., Nunberg, J. H. & Schimke, R. T. *Cell* **22**, 361–370 (1980).
40. McGrath, J. P. *et al. Nature* **304**, 501–506 (1983).
41. Amara, F. G. *et al. Nature* **298**, 240–244 (1982).
42. Waterfield, M. D. *et al. J. cell. Biochem.* **210**, 149–161 (1982).
43. Rose, J. K. & Bergmann, J. E. *Cell* **30**, 753–757 (1982).
44. Shimizu, N., Behzadian, M. A. & Shimizu, Y. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3600–3604 (1980).
45. Davies, R. L., Grosse, V. A., Kucherlapati, R. & Bothwell, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4188–4192 (1980).
46. Engelman, D. M., Henderson, R., Mchachlen, A. D. & Wallace, B. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2023–2027 (1980).
47. Noda, M. *et al. Nature* **302**, 528–532 (1983).
48. Zidovetzki, R., Yarden, Y., Schlessinger, J. & Jovin, T. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6981–6985 (1981).
49. Haigler, H. T., McKanna, J. A. & Cohen, S. *J. Cell Biol.* **81**, 382–395 (1979).
50. Schechter, Y. S., Hernaez, L., Schlessinger, J. & Cuatrecasas, P. *Nature* **278**, 835–837 (1979).
51. Schreiber, A. B. *et al. J. biol. Chem.* **258**, 846–853 (1983).
52. Schreiber, A. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 7535–7539 (1981).
53. Parker, P. J. *et al. J. biol. Chem.* (in the press).
54. Wahl, G. M., Vitto, L., Padgett, R. A. & Stark, G. R. *Molec. cell. Biol.* **2**, 308–319 (1982).
55. Wahl, G. M., Saint Vincent, B. R. & De Rose, M. L. *Nature* **307**, 516–520 (1984).
56. Barnes, D. W. *J. Cell Biol.* **93**, 1–4 (1982).
57. Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. & Mercier, R. *Nucleic Acids Res.* **9**, 43–74 (1981).
58. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
59. Gray, A., Dull, T. J. & Ullrich, A. *Nature* **303**, 722–725 (1983).
60. Benton, W. D. & Davis, R. W. *Science* **196**, 180–182 (1977).
61. Smart, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6013–6017 (1981).
62. Taylor, J. M., Illmensee, R. & Summers, J. *Biochim. biophys. Acta* **4**, 324–330 (1976).
63. Weber, W. *et al. Science* **224**, 294–297 (1984).
64. Merlino, G. T. *et al. Science* **224**, 417–419 (1984).

Conservation and change in the DNA sequences coding for alcohol dehydrogenase in sibling species of *Drosophila*

Mark Bodmer* & Michael Ashburner

Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK

The DNA sequences of the alcohol dehydrogenase (Adh) *genes of four very closely related species of* Drosophila *show that the rates of nucleotide change vary greatly in different functional domains of this gene. A phylogeny of these species based on the* Adh *sequence data is consistent with that based on polytene chromosome banding patterns.*

SPECIES of *Drosophila* are ideal for studies of DNA sequence evolution as we have very detailed knowledge of the phylogenetic relationships within many groups of the genus[1]. *Drosophila melanogaster* is an unusual member of the family Drosophilidae—it is one of a very small number of species (8 out of over 2,700) that is virtually cosmopolitan in its geographical distribution. Its closest relative, *Drosophila simulans*, is also cosmopolitan, although until recently it was rare in South-East Asia and Japan[2,3]. Recent studies of the drosophilid fauna of Africa have identified six species closely related to *D. melanogaster* and *D. simulans*[4–9]. These are *Drosophila yakuba* and *Drosophila teissieri*, widely distributed in sub-Saharan Africa; *Drosophila erecta* and *Drosophila orena*, known only from tropical West Africa; and *Drosophila mauritiana* and *Drosophila sechellia*, endemic to the islands of Mauritius and the Seychelles respectively. These eight species form the *melanogaster* species subgroup. Morphologically they are very similar, their male genitalia providing their only reliable distinguishing features. Ecologically the six Afro-tropical species may be more restricted than either *D. melanogaster* or *D. simulans*: for example *D. erecta* breeds almost exclusively on the fallen

* Present address: Celltech, 250 Bath Road, Slough SL1 4DY, Berkshire, UK.

EXHIBIT C

Reprint Series
6 September 1985, Volume 229, pp. 974–976

SCIENCE

Amplification of a Novel v-*erb*B–Related Gene in a Human Mammary Carcinoma

C. Richter King, Matthias H. Kraus, and Stuart A. Aaronson

Amplification of a Novel v-*erb*B–Related Gene in a Human Mammary Carcinoma

Abstract. *The cellular gene encoding the receptor for epidermal growth factor (EGF) has considerable homology to the oncogene of avian erythroblastosis virus. In a human mammary carcinoma, a DNA sequence was identified that is related to v-erbB but amplified in a manner that appeared to distinguish it from the gene for the EGF receptor. Molecular cloning of this DNA segment and nucleotide sequence analysis revealed the presence of two putative exons in a DNA segment whose predicted amino acid sequence was closely related to, but different from, the corresponding sequence of the erbB/EGF receptor. Moreover, this DNA segment identified a 5-kilobase transcript distinct from the transcripts of the EGF receptor gene. Thus, a new member of the tyrosine kinase proto-oncogene family has been identified on the basis of its amplification in a human mammary carcinoma.*

C. Richter King
Matthias H. Kraus
Stuart A. Aaronson
Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, Maryland 20205

The oncogenes of the acute transforming retroviruses have counterparts, designated proto-oncogenes, that are conserved within the human genome (*1*). The human *sis* proto-oncogene encodes one major polypeptide chain of platelet-derived growth factor (PDGF) (*2*), and the *erb*B proto-oncogene appears to encode the receptor for epidermal growth factor (EGF) (*3*). A number of other proto-oncogenes, like *erb*B, share nucleotide sequence homology with the tyrosine kinase–encoding *src* gene (*4*). The fact that cellular receptors for several growth factors or hormones, including the EGF receptor, possess this enzymatic activity suggests that other proto-oncogenes may encode growth factor receptors as well.

Genetic alterations affecting proto-oncogenes of the tyrosine kinase family can play a role in spontaneous tumor development. A specific translocation affecting the c-*abl* locus, for example, is associated with chronic myelogenous leukemia (*5*). Several recent studies have also documented amplification or rearrangement of the gene for the EGF receptor in certain human tumors (*6*) or tumor cell lines (*7*). We now report the detection and partial isolation of a gene that is a new member of the tyrosine kinase family and is amplified in a human mammary carcinoma. This gene is closely related to, but distinct from, the EGF receptor gene.

The identification of additional members of some proto-oncogene families has emerged from findings of related sequences amplified sufficiently in a particular tumor to allow detection (*8*). Because of our interest in genes coding for growth factor receptors, we used the v-*erb*B gene to probe for related genes that might be candidates for other receptor coding sequences. We selected moderate stringency hybridization conditions under which different oncogenes of the tyrosine family did not cross-hybridize. Thus, any gene detected might be expected to have a closer relationship to v-*erb*B than to other members of the tyrosine kinase family.

DNA prepared from tissue of a human mammary carcinoma, MAC117, showed a pattern of hybridization (Fig. 1A) that differed both from that observed with DNA of normal human placenta and from that observed with the A431 squamous-cell carcinoma line, which contains amplified EGF receptor genes (*7*). In A431 DNA, four Eco RI fragments were detected that had increased signal intensities compared to those of corresponding fragments in placenta DNA (Fig. 1A). In contrast, MAC117 DNA contained a single 6-kilobase pair (kbp) fragment, which appeared to be amplified compared to corresponding fragments observed in both A431 and placenta DNA's (Fig. 1A). These findings were consistent with the possibility that the MAC117 tumor contained an amplified DNA sequence related to, but distinct from, the cellular *erb*B proto-oncogene.

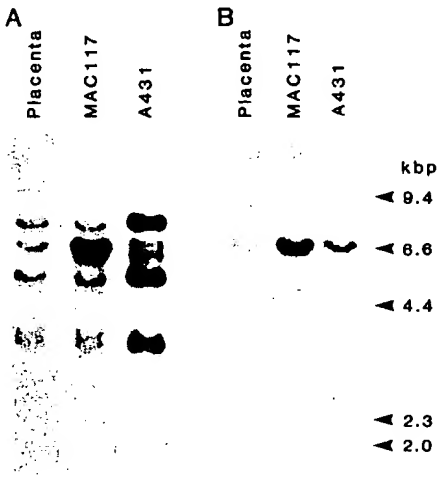


Fig. 1. Detection of v-*erb*B– and pMAC 117–specific gene fragments in normal human placenta, A431 cells, or human mammary carcinoma MAC117. DNA (15 μg) was cleaved with Eco RI, separated by electrophoresis in agarose gels, and transferred to nitrocellulose paper (*18*). Hybridization to the ^{32}P-labeled probe (*20*) was conducted in a solution of 40 percent formamide, 0.75*M* NaCl, 0.075*M* sodium citrate, at 42°C (*19*). The v-*erb*B probe (A) was a mixture of the 0.5-kbp Bam HI–Bam HI fragment and 0.5-kbp Bam HI–Eco RI fragment of avian erythroblastosis proviral DNA. The pMAC117 probe (B) was a 1-kbp Bgl I–Bam HI fragment. After hybridization, the blots were washed first in 0.3*M* NaCl plus 0.03*M* sodium citrate at room temperature, and then in 0.015*M* NaCl, 0.0015*M* sodium citrate, and 0.1 percent sodium dodecyl sulfate at 42°C (A) or at 52°C (B). Hybridization was detected by autoradiography.

To clone the 6-kbp fragment, we digested DNA from MAC117 with Eco RI, ligated it into bacteriophage λgtWES, packaged it in vitro, and transferred it to *Escherichia coli* strain BNN45 by infection. A library of 4×10^5 bacteriophages was screened by plaque hybridization with radioactive v-*erb*B DNA. Ten of 14 hybridizing phages contained a 6-kbp Eco RI fragment. Figure 2 shows the physical map of one of these phages, λMAC117, and pMAC117, a pUC12 subclone containing a 2-kbp Bam HI fragment of λMAC117 that hybridized with v-*erb*B probes. The region of pMAC117 to which v-*erb*B hybridized most intensely was flanked by Acc I and Nco I sites. Human repetitive sequences were also localized (Fig. 2, region demarcated by arrows).

By digestion of pMAC117 with Bgl I and Bam HI, it was possible to generate a single-copy probe homologous to v-*erb*B. This probe detected a 6-kb Eco RI fragment that was amplified in MAC117 DNA and possibly increased in A431 cellular DNA relative to normal DNA (Fig. 1B). The sizes of the fragments corresponded to the amplified 6-kb Eco RI fragment detected in MAC117 DNA by means of v-*erb*B (Fig. 1A). Hybridization to Southern blots containing serial dilutions of MAC117 genomic DNA indicated an approximate amplification of 5- to 10-fold when compared to human placenta DNA.

The nucleotide sequence of the portion of pMAC117 located between the Nco I and Acc I sites contained two regions of nucleotide sequence homologous to v-*erb*B separated by 122 nucleotides (Fig. 3). These regions shared 69 percent nucleotide sequence identity with both the v-*erb*B and the human EGF receptor gene. The predicted amino

induced rat neuroblastoma has been detected by DNA transfection analysis (*11*). This oncogene, designated *neu*, appears to encode a protein immunologically related to the EGF receptor (*12*). Whether the MAC117 coding sequence and *neu* represent the same or different cellular genes awaits further characterization.

Overexpression of proto-oncogenes can cause cell transformation in culture and may function in the development of human tumors. Amplification of a normal *ras* gene or its increased expression under the control of a retroviral long terminal repeat (LTR) induces transformation of NIH 3T3 cells (*13*). Expression of the normal human *sis*/PDGF-2 coding sequence in NIH 3T3 cells, which do not normally express their endogenous *sis* proto-oncogene, also leads to transformation (*14*). In Burkitt lymphoma, a chromosomal translocation involving *myc* places its normal coding sequence under the control of an immunoglobulin gene regulatory sequence (*15*). The resulting alteration in *myc* expression is likely to be causally related to tumor development (*16*). The observation of amplification of *myc* or N-*myc* in more malignant phenotypes of certain tumors has supported the idea that overexpression of these genes can contribute to the progression of such tumors (*8*, *17*). The *erb*B/EGF receptor gene is amplified or overexpressed in certain tumors or tumor cell lines (*6*). The five- to tenfold amplification of our v-*erb*B–related gene in a mammary carcinoma suggests that increased expression of this gene may have provided a selective advantage to this tumor. The isolation of a new member of the tyrosine kinase gene family amplified in a human mammary carcinoma provides an opportunity to investigate the potential role of this gene in human malignancy.

Note added in proof: Recently, Semba *et al.* (*28*) independently detected a v-*erb*B–related gene that was amplified in a human salivary gland adenocarcinoma. Nucleotide sequence analysis of this gene indicates its identity to the MAC117 gene in the regions compared.

References and Notes

1. J. M. Bishop, *Annu. Rev. Biochem.* **52**, 301 (1983); P. H. Duesberg, *Nature (London)* **304**, 219 (1983).
2. R. F. Doolittle *et al.*, *Science* **221**, 275 (1983); M. D. Waterfield *et al.*, *Nature (London)* **304**, 35 (1983).
3. J. Downward *et al.*, *Nature (London)* **307**, 521 (1984).
4. A. Hampe, I. Laprevotle, F. Galibert, L. A. Fedele, C. J. Sherr, *Cell* **30**, 775 (1982); N. Kitamura, A. Kitamura, K. Toyoshima, Y. Hirayama, M. Yoshida, *Nature (London)* **297**, 205 (1982); M. Shibuya and H. Hanafusa, *Cell* **30**, 787 (1982); T. Yamamoto *et al.*, *ibid.* **35**, 71 (1983).
5. A. de Klein *et al.*, *Nature (London)* **300**, 765 (1982); S. J. Collins and M. T. Groudine, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4813 (1983).
6. T. A. Libermann *et al.*, *Nature (London)* **313**, 144 (1985).
7. A. Ullrich *et al.*, *ibid.* **309**, 418 (1984); Y. Xu *et al.*, *ibid.*, p. 809; C. R. Lin *et al.*, *Science* **224**, 843 (1984).
8. M. Schwab *Nature (London)* **305**, 245 (1983); N. E. Kohl *et al.*, *Cell* **35**, 349 (1983).
9. R. Breathnach and P. Chambon, *Annu. Rev. Biochem.* **50**, 349 (1981).
10. J. E. Smart *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6013 (1981).
11. Alan L. Schechter *et al.*, *Nature (London)* **312**, 513 (1984).
12. J. A. Drebin, D. F. Stern, V. C. Link, R. A. Weinberg, M. I. Greene, *ibid.*, p. 545.
13. E. H. Chang, M. E. Furth, E. M. Scolnick, D. R. Lowy, *ibid.* **297**, 479 (1982).
14. A. Gazit *et al.*, *Cell* **39**, 89 (1984); M. F. Clarke *et al.*, *Nature (London)* **308**, 464 (1984).
15. R. Taub *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7837 (1982).
16. K. Nishikura *et al.*, *Science* **224**, 399 (1984).
17. M. Schwab *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4940 (1984).
18. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
19. G. M. Wahl, M. Stern, G. R. Stark, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3683 (1979).
20. P. W. J. Rigby, M. Dieckmann, C. Rhodes, P. Berg, *J. Mol. Biol.* **113**, 237 (1977).
21. A. M. Maxam and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560 (1977).
22. The following abbreviations were used for amino acids: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.
23. H. D. Lehrach, D. Diamond, J. M. Wozney, H. Boedtker, *Biochemistry* **16**, 4743 (1977).
24. A. P. Czernilofsky *et al.*, *Nature (London)* **287**, 193 (1980).
25. E. P. Reddy, M. J. Smith, A. Srinivasan, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3623 (1983).
26. A. Ullrich *et al.*, *Nature (London)* **313**, 756 (1985).
27. D. J. Lipman and W. R. Pearson, *Science* **227**, 1435 (1985).
28. K. Semba, N. Kamata, K. Toyshima, T. Yamamoto, *Proc. Natl. Acad. Sci. U.S.A.*, in press.

29. We thank I. Pastan and G. Merlino for providing the human EGF complementary DNA clone PE7 and P. di Fiore for the polyadenylated A431 RNA.

22 April 1985; accepted 3 June 1985

acid sequence of these regions was 85 percent homologous to two regions that are contiguous in the EGF receptor sequence (*7*). Furthermore, these two putative coding regions of the MAC117 sequence were each flanked by the AG and GT dinucleotides that border the exons of eukaryotic genes (*9*). These findings suggest that the sequence shown in Fig. 2 represents two exons, separated by an intron, of a gene related to the *erb*B/EGF receptor gene.

The predicted amino acid sequence of the λMAC117 putative exons is homologous to the corresponding sequences of several members of the tyrosine kinase family. The most striking homology was observed with the human EGF receptor or *erb*B (Fig. 3). In addition, we observed 42 percent to 52 percent homology with the predicted amino acid sequences of other tyrosine kinase–encoding genes. At 25 percent of the positions there was identity among all the sequences analyzed (Fig. 3). A tyrosine residue in the λMAC117 putative coding sequence, conserved among the tyrosine kinases analyzed, is the site of autophosphorylation of the *src* protein (*10*).

The availability of cloned probes of the MAC117 gene made it possible to investigate its expression in a variety of cell types. The MAC117 probe detected a single 5-kb transcript in A431 cells (Fig. 4). Under the stringent conditions of hybridization utilized, this probe did not detect any of the three RNA species recognized by EGF receptor complementary DNA. Thus, MAC117 represents a new functional gene within the tyrosine kinase family, closely related to, but distinct from, the gene encoding the EGF receptor.

There is precedent for the identification of genes related to known oncogenes on the basis of their amplification in human tumors. For example, the high degree of amplification of N-*myc* in certain malignancies made it detectable by means of the *myc* gene as a molecular probe (*8*). In the present study, a five- to tenfold amplification of a v-*erb*B–related gene in the MAC117 mammary carcinoma made it possible to identify this sequence against a complex pattern of EGF receptor gene fragments. Analysis of DNA from ten additional mammary carcinomas has not revealed amplification of the MAC117 gene. However, extensive studies will be required to determine the frequency of MAC117 gene amplification in different human malignancies.

The MAC117 coding sequence, as determined by nucleotide and predicted amino acid sequence, was most closely related to the *erb*B/EGF receptor among known members of the tyrosine kinase family. The two genes are distinct, on the basis of sequence diversity and transcript size. Detailed structural analysis of the complete coding sequence should give insights into the possible functions of v-*erb*B–related gene. Nevertheless, because of its close relationship to the sequence of the EGF receptor, it is possible to speculate that the MAC117 coding sequence may also be derived from a gene encoding a growth factor receptor. An oncogene in a chemically

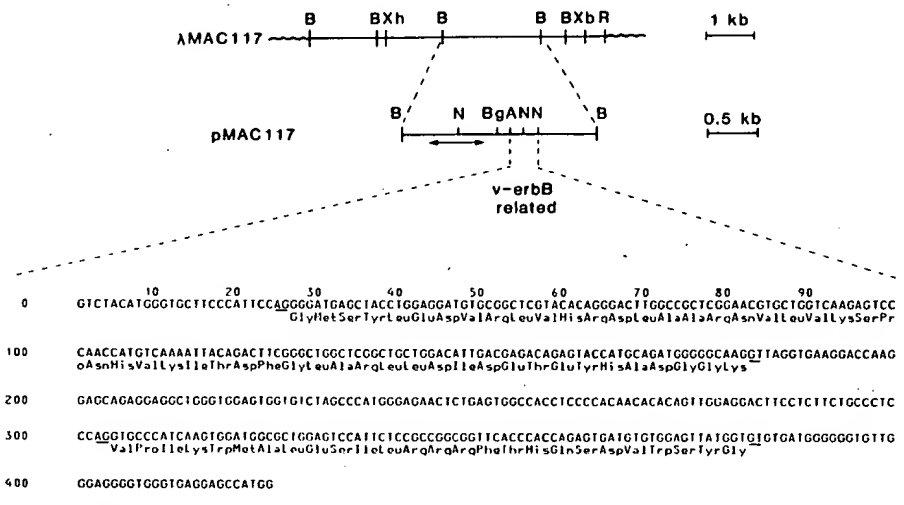


Fig. 2. Restriction-site map of λMAC117 and plasmid pMAC117. A, Acc I; B, Bam HI; Bg, Bgl I; N, Nco I; R, Eco RI; X, Xba I; Xh, Xho I. The sites were located by electrophoretic analysis of the products of single and double digestion. Regions homologous to v-*erb*B or human repetitive sequences (region flanked by arrows) were located by Southern blot hybridization (*18*) with the v-*erb*B probe or total human DNA made radioactive by nick translation (*20*). Hybridization conditions were as described in Fig. 1A. The nucleotide sequence of pMAC117 between the Acc I site and the Nco I sites and regions of encoded amino acid sequence homologous to the EGF receptor are shown. The AG or GT dinucleotides flanking the putative coding regions are underlined. To determine the sequence, Nco I, Hinf I, and Sau 96 I fragments were labeled at the 3′ termini by means of the large fragment of *E. coli* DNA polymerase, separated into single strands by gel electrophoresis, and chemically degraded (*21*).

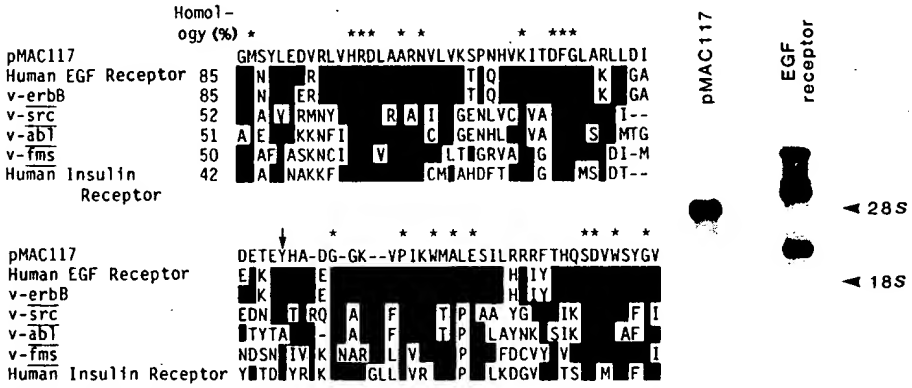


Fig. 3 (left). Comparison of the putative encoded amino acid sequence in pMAC117 with known tyrosine kinase sequences. Black regions represent homologous amino acids. Differing amino acid residues are shown in one-letter code (*22*). Amino acid positions conserved in all sequences are denoted by *. The tyrosine homologous to that autophosphorylated by the v-*src* protein (*10*) is shown by an arrow. The v-*abl* sequence contains a tyrosine residue in this region displaced by two positions. The amino acid sequences of human EGF receptor (*7*), v-*erb*B (*4*), v-*src* (*23*), v-*abl* (*24*), v-*fms* (*4*), and human insulin receptor (*25*) were aligned by the computer program described (*26*). The homology observed with the predicted amino acid sequences of v-*yes* and v-*fes* was 51 percent and 48 percent, respectively. Fig. 4 (right). Detection of distinct messenger RNA species derived from the λMAC117 gene and the human EGF receptor gene. Polyadenylated messenger RNA of A431 cells was separated by denaturing gel electrophoresis in formaldehyde (*23*), transferred to nitrocellulose (*18*), and hybridized under stringent conditions (50 percent formamide, 0.075*M* NaCl, 0.75*M* sodium citrate, at 42°C) with ^{32}P-labeled probe from pMAC117 (Bgl I–Bam HI fragment) or human EGF receptor complementary DNA (PE7: 2-kb Cla I inserted fragment). Filters were washed under conditions of high stringency (0.015*M* NaCl plus 0.0015*M* sodium citrate at 55°C). Hybridization was detected by autoradiography with exposure times of 4 hours for the pMAC117 probe and 1 hour for the human EGF receptor probe.